National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

Spatial Reuse in Broadcast Networks

BY

Jacek Królikowski

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science.

DEPARTMENT OF COMPUTING SCIENCE

Edmonton, Alberta
Fall 1994

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-95051-X

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Jacek Królikowski

TITLE OF THESIS: Spatial Reuse in Broadcast Networks

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: 1994

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

(Signed) J. Królikowski

Jacek Królikowski
#103, 10511-92 Street
Edmonton, Alberta
T5H 4E6

Date: July 11/94

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Spatial Reuse in Broadcast Networks** submitted by Jacek Królikowski in partial fulfillment of the requirements for the degree of Master of Science.

. . . . . .
Dr. W. Dobosiewicz (Supervisor)

. . . . . .
Dr. P. Gburzyński (Co-supervisor)

. . . . . .
Dr. W. Rozmus (External)

. . . . . .
Prof. U. Maydell (Examiner)

Dr. T.A. Marsland (Chair)

Date: July 11/94

To my Mother

Abstract

This thesis discusses the performance of the chordal bus and chordal ring architecture used with three medium-access-control protocols: Simple Reuse Protocol using Previous Slot Information, Optimized Reuse Protocol using Previous Slot Information and Meta Protocol. The thesis also presents three starvation-prevention mechanisms: SAT—Global Fairness Mechanism, LFM—Local Fairness Mechanism and MLFM—Modified Local Fairness Mechanism.

The Optimized Reuse Protocol is a *capacity-1* protocol; its maximum throughput does not depend on the network size and the transmission rate. The protocol allows conventionally built stations partial, software-supported destination release significantly improving the aggregated network throughput.

The Modified Local Fairness Mechanism dynamically divides the transmission medium into local communities of interest and performs better than all known global fairness mechanisms.

To investigate the performance of new protocols, simulation models are used. The SMURPH simulation package is used to develop the simulators.

Acknowledgements

I would like to take this opportunity to express my sincere thanks to those who helped me throughout the course of this study. I am particularly grateful to my supervisor Dr. Włodek Dobosiewicz for his assistance, support and personal concern for myself and the study. His expectations for maintaining steady progress and producing quality work have been driving forces for me. I have benefited from his knowledge, experience and willingness to discuss many different aspects of the project.

A sincere appreciation is expressed to my co-supervisor Dr. Paweł Gburzyński for his technical assistance and support for this study. His insight into the area of the network protocol simulation was the source of many fruitful discussions.

I am also grateful to the members of the examine committee, Prof. U. M. Maydell, Dr. W. Rozmus for the time they took to read and comment on this thesis.

Contents

1 Introduction **1**

1.1 Motivation 2

1.2 Thesis Overview 2

1.3 Related Work 3

1.3.1 Bus Architecture 3

1.3.2 Ring Architecture 4

1.4 Performance criteria 4

2 Architecture Overview **7**

2.1 Architecture 7

2.1.1 Regular station 7

2.1.2 Erasure node 9

2.1.3 Slot generator 9

2.1.4 Links 10

2.2 Enhancements 11

2.2.1 Wiring closet switch 11

2.2.2 Meta station 13

3 Architecture: details **14**

3.1 Bus Architecture 14

3.1.1 Plain architecture 14

3.1.2 Chordal architecture 16

3.2 Ring Architecture 17
3.2.1 Plain architecture 19
3.2.2 Chordal architecture 20

4 Mac protocols **22**
4.1 Protocols for the bus architecture 23
4.1.1 Simple 25
4.1.2 DQDB 26
4.1.3 PSI 27
4.2 Protocols for ring architecture 29
4.2.1 FDDI 29
4.2.2 Metaring 31
4.2.3 CSMA/RN Ring 32
4.3 Starvation Prevention Mechanisms 33
4.3.1 SAT mechanism 34
4.3.2 Local Fairness Mechanism 35
4.3.3 Modified Local Fairness Mechanism 39

5 Simulation Environment **42**
5.1 The Simulator 42
5.2 Performance Measurements 43
5.3 Traffic Conditions 44

6 Simulation results **45**
6.1 Plain architectures 46
6.1.1 Bus Architecture 46
6.1.2 Ring Architecture 50
6.2 Chordal architectures 53
6.2.1 Bus Architecture 53
6.2.2 Ring Architecture 56
6.3 Comparison 58

7 Conclusions and Future Work 63

Bibliography 66

A Dual Chordal Bus 71

B Dual Chordal Ring 73

List of Figures

2.1 Regular station structure. 8
2.2 Active tap configuration. 8
2.3 Erasure node configuration. 10
2.4 Typical usage of a wiring closet. 12
2.5 Wiring closet switch. 12
2.6 Meta Station. 13

3.7 General dual bus architecture. 15
3.8 15
3.9 16
3.10 Metabus versus bus with erasure nodes. 17
3.11 A modified bus architecture. 18
3.12 Dual ring architecture. 19
3.13 Chordal ring architecture. 21

4.14 The slot structure. 24
4.15 PSI slot structure. 27
4.16 FDDI frame format. 30
4.17 Starvation in METARING. 33
4.18 SAT-based Global Fairness Mechanism (GFM). 34
4.19 Local starvation on a bus. 35
4.20 LFM on a bus. 36
4.21 LFM state transition diagram (bus architecture). 38

4.22 Local starvation on a bus. General case. 40

6.23 RBus performance. 47
6.24 Placement of Erasure Nodes. 47
6.25 ROptBus performance. 48
6.26 MetaBus performance. 49
6.27 RRing performance. 51
6.28 ROptRing performance. 51
6.29 METARING performance. 52
6.30 Placement of Wiring Switch Nodes in Chordal Bus. 53
6.31 RSBus performance. 54
6.32 RSOptBus performance. 55
6.33 MetaSBus performance. 56
6.34 RSRing performance. 57
6.35 RSOptRing performance. 58
6.36 S METARING performance. 59
6.37 Performance of Plain Bus Architecture. 60
6.38 Performance of Chordal Bus Architecture. 61
6.39 Performance of Plain Ring Architecture. 61
6.40 Performance of Chordal Ring Architecture. 62

A.41 71
A.42 72

B.43 73
B.44 74

List of Symbols

ACF Access Control Field

AMD Absolute Message Delay

ATM Asynchronous Transfer Mode

BWB Bandwidth Balancing

CBRMA++/SR Cyclic-Balanced Reservation Multiple Access ++ with Slot Reuse

CRMA Cyclic Reservation Multiple Access

CSMA/RN Carrier Sense Multiple Access / Ring Network

DAS Double Attach Stations

DQDB Distributed Queue Dual Bus

FDDI Fiber Distributed Data Interface

GFM Global Fairness Mechanism

LAN Local Area Network

LFM Local Fairness Mechanism

MAC Medium Access Control

MAN Metropolitan Area Network

MAT Message Access Time

MLFM Modified Local Fairness Mechanism.

MMAT Mean Message Access Time

ORP/PSI Optimized Reuse Protocol using Previous Slot Information

PSI Previous Slot Information

SAS Single Attach Stations

SAT Synchronous Allocation Time

SRP/PSI Simple Reuse Protocol using Previous Slot Information

THT Token Holding Timer

TRT Token Rotation Time

TTRT Target Token Rotation Time

WAN Wide Area Networks

Chapter 1

Introduction

The tremendous development of the computer technology in the past ten years has increased the demand for networks with a high throughput. Many applications like multi media, tele conferencing, digital medical imaging, hyper–text with hyper–links require sending large amounts of data in a very short time. The introduction of X–terminals with large sizes of computer screens (in pixels), disk less workstations, and network file systems has caused noticeable increase in the load of local computer networks. Although some relief can be achieved by partitioning the network with the help of fast routers and switches, it is clear that fiber optics technology with new network protocols will be needed in the near future to cope with the increased demands.

The request for high speed networks is not limited to LANs. Even higher throughput is expected from *Metropolitan Area Networks* (MAN) and *Wide Area Networks* (WAN).

One of the parameters which facilitate the description of the network span is the parameter "a" [Kle75] defined as the ratio of the network–wide propagation delay to the time required to transmit a packet. Clearly, when the length of the network increases or network devices are faster, the value of a increases as well. Early protocols like CSMA/CD [Eth80, IEEa] and TOKEN RING [IEEb] which allowed only one packet in the medium at a time, performed well when a was smaller than 1. Even modern

protocols using fiber optic links like FDDI do not perform well when a is greater than 1. In gigabit networks the value of a is usually greater than 10. Clearly new protocols able to deal with large a's are needed for contemporary distributed computer systems.

1.1 Motivation

Even though *Asynchronous Transfer Mode* (ATM) [CCI91a, CCI91c, CCI91b] is seen by some researchers as a protocol which will solve all problems of networks of the future, some other authors are not so enthusiastic [CDGK92]. ATM is still in its infancy and many problems of it are not solved yet.

In areas where a fast and reliable data transfer is needed, more data transfer oriented protocols are necessary. Investigating properties of such very high speed protocols is the subject of this thesis.

1.2 Thesis Overview

The thesis is divided into seven chapters.

Chapter 1 gives a brief introduction to the problems encountered in very high speed protocol design and the current research within these systems. The discussion concentrates on bus and ring shaped networks. This chapter discusses some of the performance criteria used in the evaluation of network protocols.

Chapter 2 describes the building blocks used in such protocols as FDDI and DQDB and some new devices needed by novel protocols as METARING and *Reuse* protocols.

Chapter 3 presents the elements of a new chordal architecture which is introduced in this thesis.

Chapter 4 describes MAC-level protocols investigated in the thesis. Among others, it presents in details a new *Optimized Reuse Protocol using Previous Slot*

Information protocol. This chapter also presents some starvation–prevention methods used by the investigated protocols. Among them, a new method—*Modified Local Fairness Mechanism* is presented.

Chapter 5 describes important properties of the SMURPH simulation package and the simulation environment supplied by it. This package was the main tool used to write simulators of the protocols investigated in this thesis.

Chapter 6 presents simulation results of the protocols investigated in the thesis. The simulation results are shown both in a graphical and numerical form.

Chapter 7 presents conclusions and future research directions.

1.3 Related Work

The research in packet and slotted protocols for optical networks focuses mainly in the areas: bus, ring, and mesh networks. Each of these architectures have different properties and require different approaches in analyzing them. Bus–shaped networks are potentially unfair, ring–shaped networks are fair but sometimes starvation prone, mesh networks with deflection routing rearrange incoming slots and can have unbound message delays. Since mesh networks are fundamentally different from bus and ring networks they are not described in this thesis. A good introduction to these types of networks and their properties can be found in [Max85].

1.3.1 Bus Architecture

One of the best known protocols for the bus architecture is undoubtedly DQDB. A short description of this protocol can be found in Section 4.1. Because of its unfairness, which is exhibited when the network is close to saturation point, some improvements to DQDB have been proposed. One of the countermeasures is *Bandwidth Balancing* (BWB) which only works after a long reaction time [vA90]. Another well known protocol—CRMA I and its successor CRMA II developed in IBM Zurich—is described

in [Nas90, vALZ92]. It performs better than DQDB under heavy load conditions. CBRMA++/SR [BDG93] uses cyclic balanced reservation and slot reuse to achieve excellent performance under mixed traffic conditions. SIMPLE [Lim90] belongs to the simplest yet effective protocols for bus architecture and is described in Section 4.1.

1.3.2 Ring Architecture

Astonishingly, the ring architecture is not as popular among the researchers as the bus architecture. Nevertheless, many new protocols have been proposed for this interesting architecture. FDDI [FDD90] belongs to the most successful (commercially) protocols for this architecture; it is described in the Section 4.2 although its design limits FDDI application only to medium speed and size networks. *Weak token* protocols [DG92b, DG92a] relax the strict FDDI requirement that a station can transmit its packet only if it holds a token, and perform much better for very fast ring networks. METARING [CO89b] (see Section 4.2) is another MAC protocol proposed for optical ring networks. It uses hardware–supported *destination release* to achieve unrivalled performance.

1.4 Performance criteria

To be able to compare the performance of several investigated systems common performance measures must be defined. The performance of a network protocol can be described by many measures like throughput, absolute message delay, absolute packet delay, message access time, packet access time. A message generated at a station must sometimes be divided into many packets due to the requirements of the network protocol (for example fixed size of slots in a slotted protocol), hence different measures are needed for messages and packets.

Throughput T of the network is equal to the **number of bits received per unit of time**, as in bits per second [bits/sec]. This measure however does not describe the performance of MAC protocols properly.

Throughput T of the network is determined by following factors:

- Transmission rate of the network equipment (γ).
- MAC protocol.
- Arrival rate of packets to the whole network (ρ), expressed in [bits/sec].

In general a high arrival rate can exceed the ability of the MAC protocol to handle arriving packets. Such a state is unacceptable from a user point of view, and this state must be avoided. For this reason we are interested in the state, in which the MAC protocol is able to service all arriving packets. In such state $T = \rho$.

We introduce also a notion of *Maximum Throughput* T_{max} as a maximum T such that $T = \rho$. For higher values of ρ the network cannot sustain a throughput of ρ which induces unbound delays.

Since we are interested in the performance of MAC protocol we would like to abstract from differences in transmission rate γ. We achieve that by dividing every time-related quantity by the transmission rate; after this division, throughput, instead of being expressed in [bits/sec] becomes a unitless-quantity. More specifically it is equal to the number of bits received in the network in the amount of time needed to transmit 1 bit.

The ratio of the payload to the size of a packet, since packets sent by station consist not only of payload bits but also of other bits, like header trailer and inter–slot gaps, must also be considered. The *Normalized Throughput* $\mathcal{T}$ thus equals:

$$\mathcal{T} = \frac{T}{\gamma} \frac{N_{payload}}{N_{payload} + N_{header} + N_{gaps}}$$

where:

$N_{payload}$ – the number of payload bits in a slot

N_{header} – the number of header bits in a slot

N_{gaps} – the sum of all gaps in a slot

For instance in the ATM protocol, where $N_{payload} = 48$, $N_{header} = 5$ and $N_{gaps} = 0$ (no inter-cell gaps) the throughput is $\mathcal{T} = \frac{T}{\gamma}\frac{48}{53} = 0.906\ \frac{T}{\gamma}$

The maximum normalized throughput thus is:

$$\mathcal{T}_{max} = \frac{T_{max}}{\gamma}\ \frac{N_{payload}}{N_{payload} + N_{header} + N_{gaps}}$$

Since $\mathcal{T}_{max}$ can be found only after an infinite amount of time, we introduce a notion of $\mathcal{T}_{real_max}$ defined as a throughput such that the maximum message access time reaches 10,000 slots. We assume that higher maximum message access time indicates that the MAC protocol cannot service all arriving packets.

In the rest of the thesis the "throughput" is meant to denote "normalized throughput" $\mathcal{T}$ unless explicitly expressed in [bits/sec].

The *message access time* (MAT) is the amount of time passed since the message was queued at the sender to the moment the last packet of the message has been successfully transmitted by the sender.

The *absolute message delay* (AMD) is the amount of time passed since the message was queued at the sender station to the moment the last packet of the message has been received at its destination. It means that AMD consists of the *message access time* and the time required by the last packet to reach its destination.

To avoid blurring the results by taking into account message segmentation and reassembly, the message size is exactly equal the packet size. This simplification is justified, since adopting other message sizes would influence all the protocols and topologies in an identical way.

Chapter 2

Architecture Overview

2.1 Architecture

Both ring and bus shaped networks (using optical fiber as a transmission medium) are made up of regular and erasure stations connected to links. In the bus architecture used with slotted protocols, two stations, called *slot generators*, are needed. Connection between stations and the fiber optic link is usually done with the help of passive or active taps. The structure and implementation of all their elements is presented in this section. Some of the details—especially in the area of fiber optics, are beyond the scope of this thesis and are omitted.

2.1.1 Regular station

A regular station collects local traffic addressed to it from the network and inserts its own packets into the network. In the slotted model, the station can only use empty incoming slots for this purpose. It also receives filled slots addressed to itself. It cannot empty these slots because of the principles used to build taps. The structure of the regular station is presented in Figure 2.1.

The construction of a tap allows the station to receive and send information at practically the same time. The physical construction of a tap differs from producer to producer, but generally one can recognize two types of taps: passive and active.

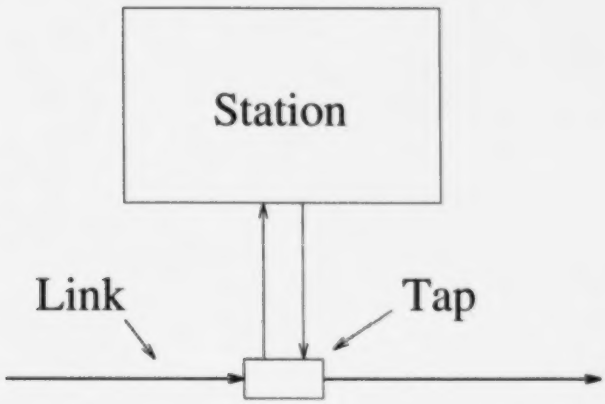


Figure 2.1: Regular station structure.

Passive taps introduce no latency. On the other hand, active taps introduce a small amount of delay, but it usually is so small, that it can be ignored in a network model (i.e., assumed to be subsumed by the propagation delay).

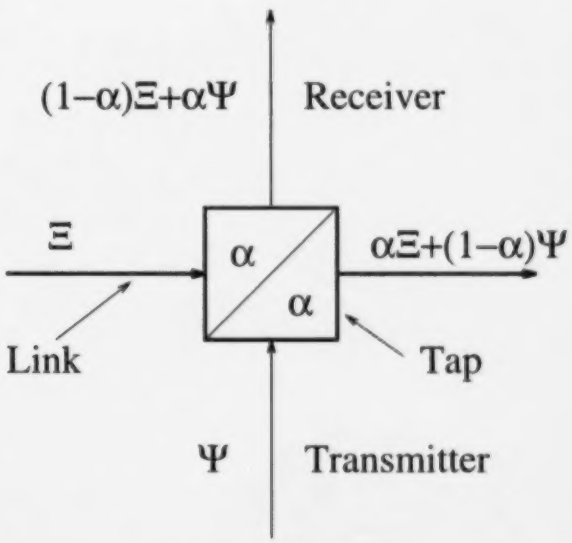


Figure 2.2: Active tap configuration.

Figure 2.2 presents the construction of an active tap. The network interface logic can change the coefficient α which regulates how much light passes through the tap. If α equals 1 then the output is directly connected to the input link of the tap. If α is equal 0 the input signal is blocked and the station can insert its signal Ψ while at

the same time receiving the input signal. This mode of operation is especially useful if the station decides to set a symbol in the header of the slot as it is often required when claiming an empty incoming slot.

2.1.2 Erasure node

An erasure node plays a special role in a networking system. Because a regular station cannot remove received packets from the network, two solutions are possible. One possibility is to do nothing. This approach is used in such protocols as FDDI and DQDB. In the case of FDDI the received packets are removed automatically by the token holding station. The DQDB protocol allows the received packets to travel through the remaining portion of the bus and to be discarded automatically at its end. The other approach is to introduce stations, which by the changes in the construction of the tap, are capable to remove all received packets (slots) from the network. Because the station must read the whole header to make a decision about the destiny of the incoming packet a buffering of packets (their headers) is necessary. A *cut-through* buffer allows simultaneous receiving and sending of the same slot. Figure 2.3 shows the layout of such a node.

As can be seen, the construction of such an interface is more complicated than in the case of a regular station. An insertion buffer allows the station to receive a packet header or even a whole incoming packet without re-sending it. This way the station can decide if the packet should be retransmitted or removed from the network. Obviously, erasure nodes introduce delays in the network. Fortunately, the amount of delay is constant (per erasure node) and limited (at least in principle) to the size of the packet's header.

2.1.3 Slot generator

The simplest active element in a bus network is the slot generator. Its only purpose is to insert empty packets into the network. It can be implemented as a stand-alone

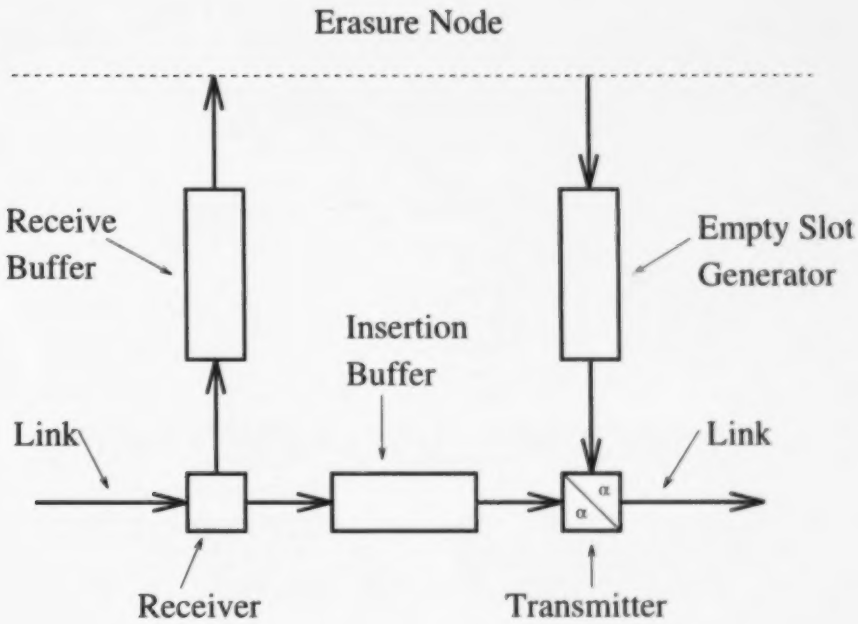


Figure 2.3: Erasure node configuration.

device or the last station attached to the end of the bus can perform this function in addition to its other activities.

Some of the slot generators can also gather some traffic statistics used to maintain the network, and recognize some error conditions occurring in the network (e.g. lost and corrupted packets).

2.1.4 Links

An optical fiber link consists of a single glass fiber contained within a protective coating. The fiber consists of two parts: the glass core and a glass cladding with a lower reflective index. Depending on the construction and the diameter of the fiber, optical links can be divided into multi-mode and mono-mode ones. Because the mono-mode optical fibers can operate at the rate of many gigabits per second, they are used to build very high speed networks.

Links are represented as point-to-point communication channels connecting pairs of adjacent nodes. Unidirectional links were chosen to represent fiber-optic connec-

tions, in our simulation models of bus and ring networks (described in Section 5.1), all links have the same physical properties [1], i.e., the transmission capacity is the same for all links in the entire network. All distances between nodes (and lengths of links) are expressed in bits.

2.2 Enhancements

The networks' elements described above can be used to build traditionally shaped networks like rings (including double rings) and busses. To construct less traditional architectures like chordal rings and folded busses new elements are needed. To these new elements one can count the *wiring closet switch* and a station allowing hardware–supported *destination release*. Such a station is called here a *meta station*.

2.2.1 Wiring closet switch

Although the virtual network architectures commonly used to analyze network protocols are usually very regular, the shape of real networks is far from perfect. One of the elements used in the construction of real networks is the wiring closet. This *passive* device is used to facilitate the wiring process and is usually a place where many fragments of the network can be accessed. The typical situation encountered in the wiring process is presented in Figure 2.4.

One of the new network elements proposed in this thesis is the *wiring closet switch*. Unlike the wiring closet described above, this network element is *active* and plays a significant role in the communication system. It is used not only to connect a collection of links into the desired architecture but also, by routing the incoming slots, it changes the way the slots travel to reach their destination. The wiring closet switch also plays the role of an erasure node. A simplified construction of such a node is presented in Figure 2.5.

[1] but, of course, they can be of different lengths.

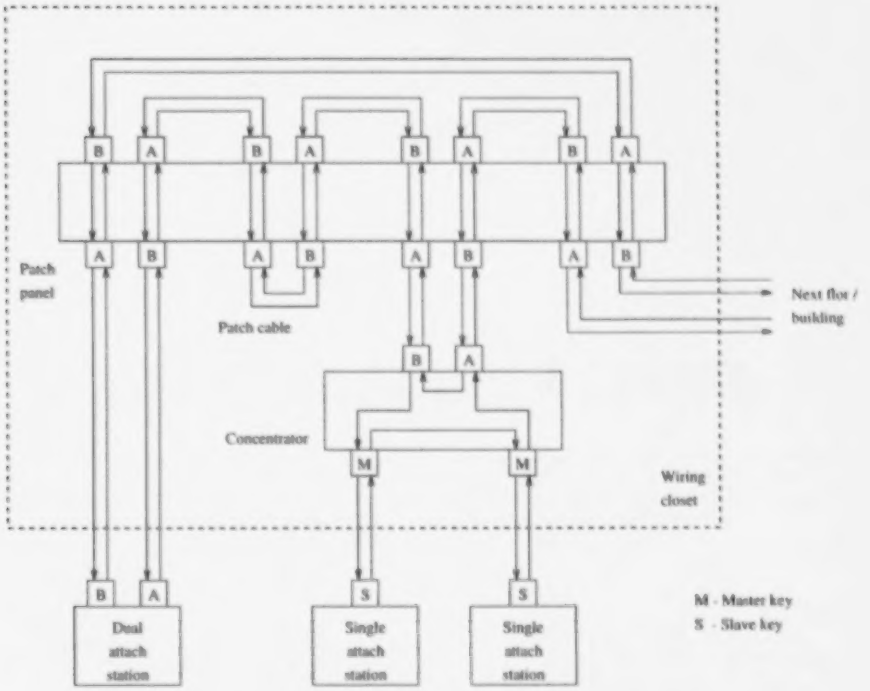


Figure 2.4: Typical usage of a wiring closet.

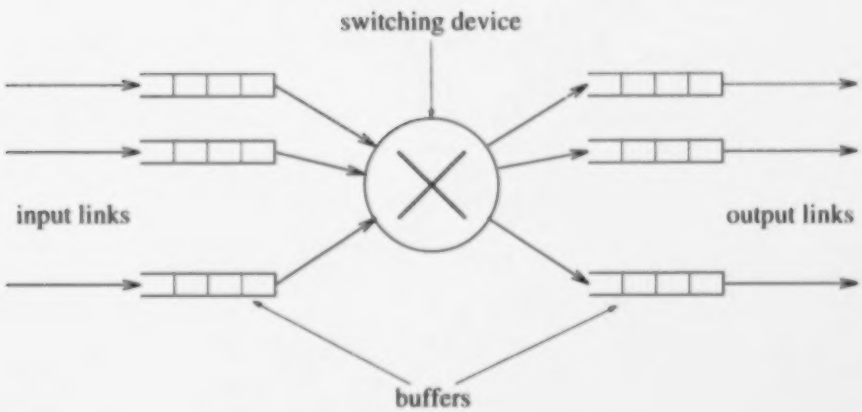


Figure 2.5: Wiring closet switch.

The number of incoming and outgoing links for a wiring closet switch can be arbitrary, but in this thesis wiring closet switches with two incoming and two outgoing links are used to build chordal busses and rings.

2.2.2 Meta station

Erasure nodes typically do not originate any traffic. If an erasure node itself is a traffic source, it is called a *meta station*. In such a case, the structure of the erasure node must be augmented by *receive* and *transmit buffers*. The *receiver* and *transmitter* may be implemented as a simple *multiplexer* and *de-multiplexer*. Figure 2.6 presents the structure of a meta station.

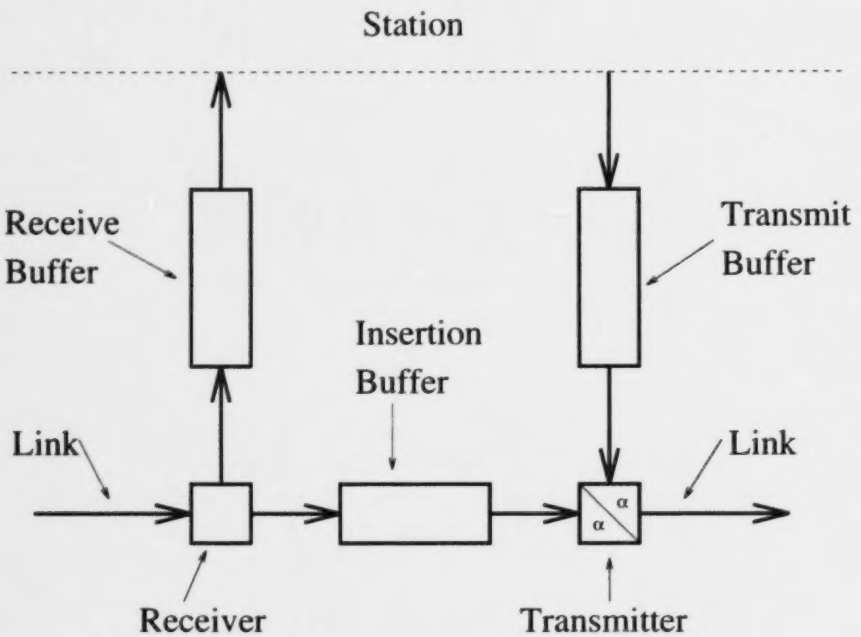


Figure 2.6: Meta Station.

The construction of the meta station enables the hardware–supported destination release. This property, when coupled with an appropriate MAC–level protocol, greatly improves the network throughput eliminating the bandwidth wasted by forwarding packets already received.

Chapter 3

Architecture: details

This chapter describes the details of the dual bus and ring architectures. In both cases two distinct architecture concepts are described:

- Plain architecture
- Chordal architecture

3.1 Bus Architecture

3.1.1 Plain architecture

The dual bus architecture (see Figure 3.7) can deliver the maximum normalized throughput of 2.0 when used with such a protocol as DQDB [DQD91], SIMPLE [Lim90], or CRMA I [MNW90]. Only two packets can be inserted into the network at a time (one in each direction), hence the maximum throughput of two.

If meta stations are used instead of regular ones the maximum throughput increases to 4.0.

To prove this, assume that there is a bus of length $\mathcal{L} = 1$ (see Figure 3.8) with meta stations evenly distributed on it. The bus represents one of the two busses of the dual-bus architecture. Each station in the bus sends v packets in a period of

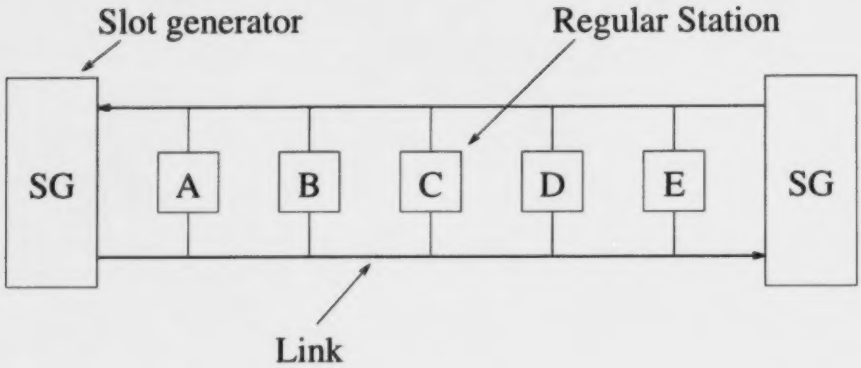


Figure 3.7: General dual bus architecture.

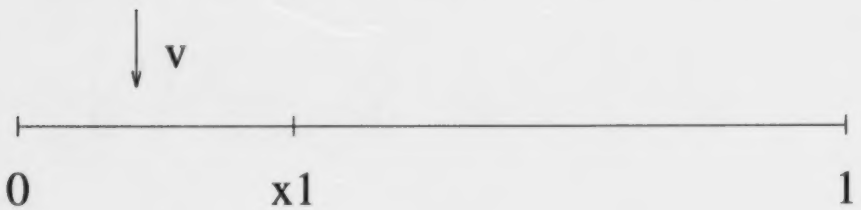


Figure 3.8:

time. All stations in the bus have the same probability of receiving packets, i.e., the traffic pattern is uniform.

Now assume that there is a point x_1 on the bus. The load at point x_1 can be calculated as the number of packets sent on the segment $[0, x_1]$ minus the number of packets received in this segment.

We focus on the forward bus (i.e., with packets traveling towards 1). Since a station located at $s \in [0,1]$ inserts $v(1-s)$ packets into this bus and vs packets into the other bus, the number of sent packets on the segment $[0, x_1]$ can be calculated as:

$$\int_0^{x_1} v(1-x)\,dx = vx_1(1-\frac{x_1}{2})$$

To calculate the number of received packets assume that there are two bus segments dx and dy as presented in Figure 3.9. Stations in segment dx send packets to stations in segment dy with probability $\frac{dy}{1-x}$. The number of received packets in

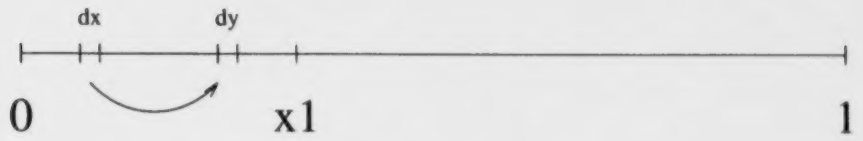


Figure 3.9:

segment $[0, x_1]$ can be calculated as:

$$\int_0^{x_1} \int_x^{x_1} v(1-x)dx\frac{dy}{1-x} = \frac{1}{2}v{x_1}^2$$

Now, the load in any arbitrary point x in the segment [0, 1] can be calculated as:

$$vx(1-\frac{x}{2}) - \frac{1}{2}vx^2 = vx(1-x)$$

The load cannot exceed the capacity of the bus (which is normalized to 1). This load function reaches its maximum at $x = 1/2$, where it is equal to 1 for $v = 4$. □

On the other hand, if a combination of regular and erasure nodes is used the maximum throughput falls somehow between 2.0 and 4.0. The value depends on the number and the relative position of the erasure nodes. It is easy to see that the throughput of the network with erasure nodes put in between plain stations will never reach the throughput of the *metabus* (a bus with all stations being meta stations).

Figure 3.10 presents a key difference between the metabus and a bus with erasure nodes. In the system with erasure nodes a station cannot reuse any incoming slots addressed to itself. Only the next station can do so. In the case of the metabus architecture the freed slots can be reused immediately. This phenomenon causes a slight increase in the maximum throughput of the latter system.

3.1.2 Chordal architecture

The introduction of wiring closet switches helps a bus network to achieve better performance. Figure 3.11 presents such a system. As in the case of the *chordal ring* (which will be introduced formally in Section 3.2.2) the existence of wiring closets inspired this solution. Unlike in the case of a chordal ring, the modified bus architecture increases the asymmetry in the bus system.

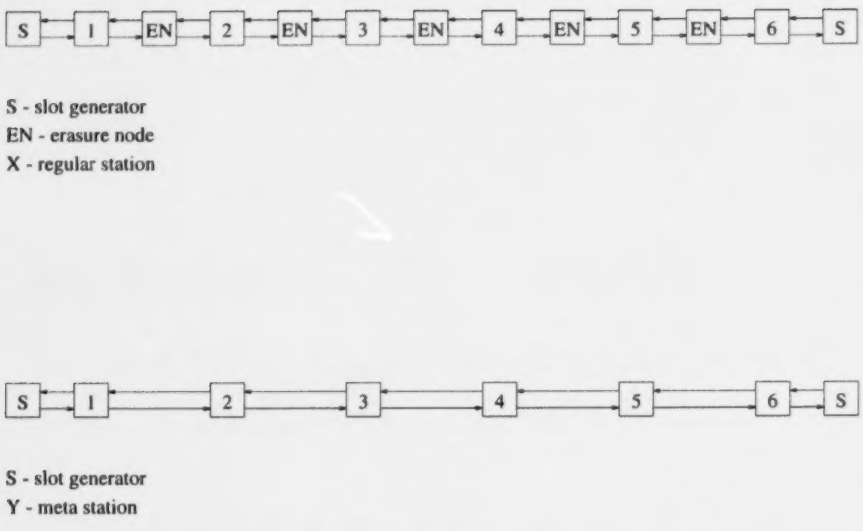


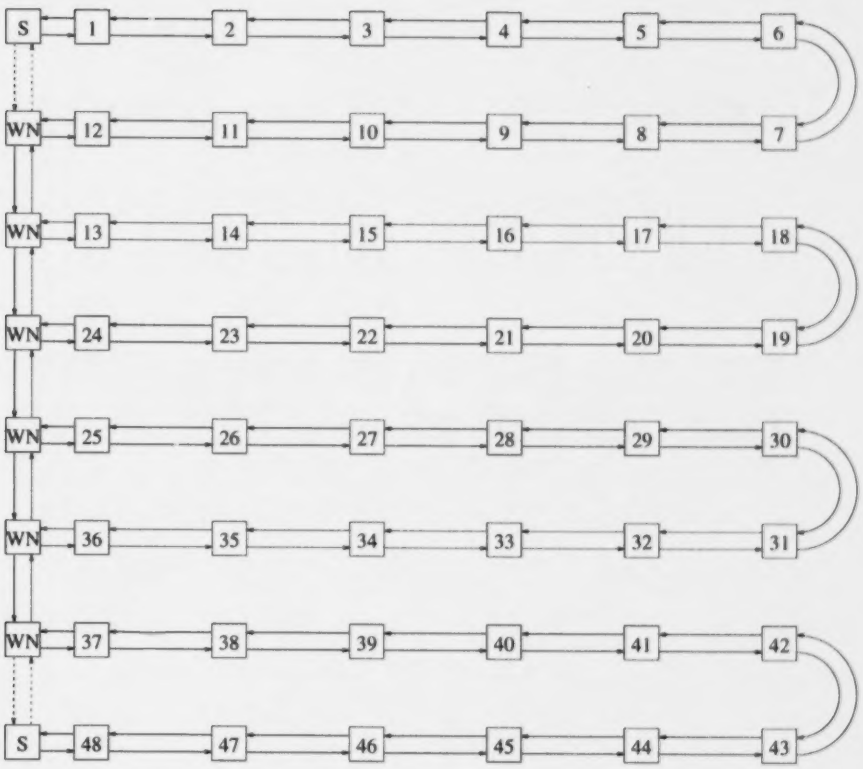
Figure 3.10: Metabus versus bus with erasure nodes.

To achieve a balanced load in different sections of the new system, the wiring closet switches must be placed properly. The load at the upstream end of the bus is much higher than at the other end. By moving the first wiring closet switch closer to the upstream end, one can achieve an even distribution of the load in the system.

3.2 Ring Architecture

The ring is a very promising architecture for very high speed networks. Unlike the bus architecture, which by its inherited asymmetry introduces problems with fairness, the ring architecture is symmetric and unless the MAC protocol implemented for it introduces asymmetry, all stations have equal access to the transmission medium.

It is rather surprising that although FDDI exists as a successful product for medium speed networks (100 Mbps), few protocols have been proposed for gigabit ring networks.



S - slot generator
WN - wiring closet switch
X - station

Figure 3.11: A modified bus architecture.

3.2.1 Plain architecture

A ring architecture consists of one, two or several fiber optic rings and a number of stations connected to them. Depending on the architecture, a station can be connected to one, two or all rings. Figure 3.12 presents a dual counter–rotating ring with *Single Attach Stations* (SAS) and *Double Attach Stations* (DAS). Such an architecture is primarily used by FDDI.

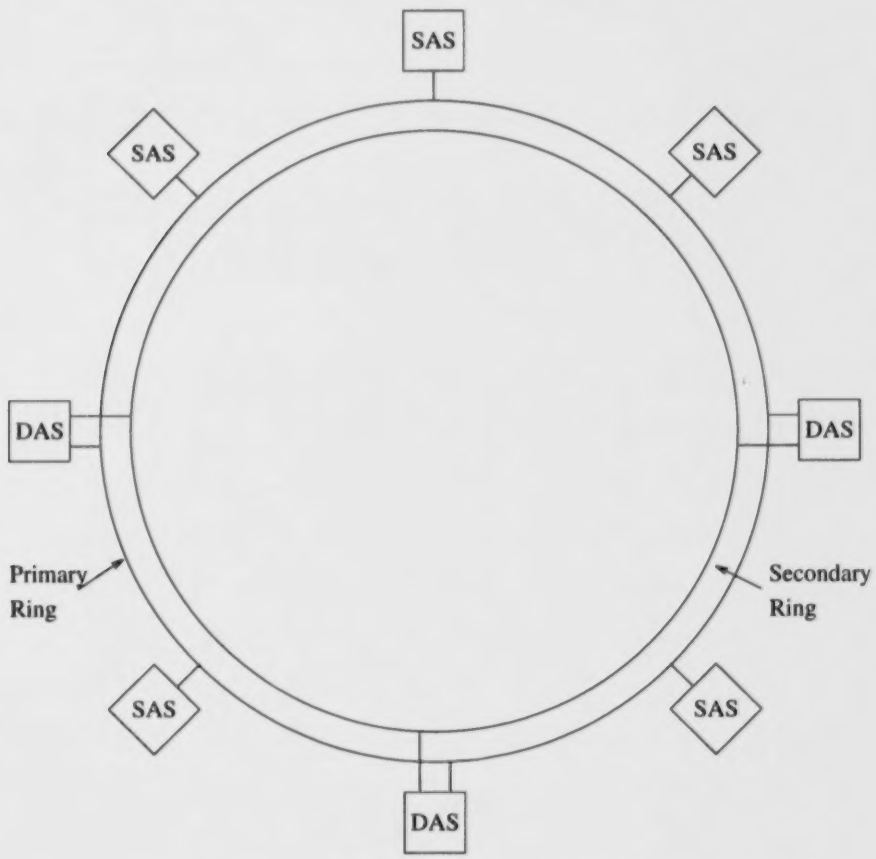


Figure 3.12: Dual ring architecture.

If only the primary ring is used for information exchange (see Figure 3.12), and the secondary ring is only used as backup, the maximum throughput theoretically achievable by this architecture is 1.0. This is what FDDI yields if it is tuned for maximum throughput. If two rings were used, FDDI would yield the maximum throughput

of 2.0.

It is easy to show that the maximum theoretical throughput of a double counter-rotating ring with double attach stations, assuming uniform load in all stations, is 8.0. If a station wants to send a packet to another station on the ring it can choose one of two rings. Because the rings are counter-rotating, there are two distances between any pair of stations. If a station always chooses the shorter distance to the destination station, the maximum distance a packet must travel is half a length of the ring. Thus the average distance a packet must travel in such a network is 1/4 of the length of a ring and on average eight stations can insert their packets into the rings. To achieve this maximum throughput, received packets must be removed by the destinations. Otherwise some of the bandwidth would be wasted. The process of removing packets by the destination station is called *destination release* or *spatial reuse*.

A good example of a protocol based on two counter-rotating rings and destination release is METARING [CO89b]. The maximum throughput achieved by this protocol is about 7.6 with mean message access time reaching 1,000 slots[1].

3.2.2 Chordal architecture

A higher maximum throughput for the ring architecture can be achieved by adding new elements to the network. One of such elements is the *wiring closet switch* described in Chapter 2.2.1. The new architecture is known under the name *chordal ring* [AL81], as illustrated in figure 3.13, which shows a chordal ring architecture with eight wiring closet switches. It should be noticed that the length of chordal links in the simulation model used in this thesis is much shorter than that shown on Figure 3.13 and is equal to one half of the slot length.

[1]The notion of the mean message access time is defined in Section 1.4.

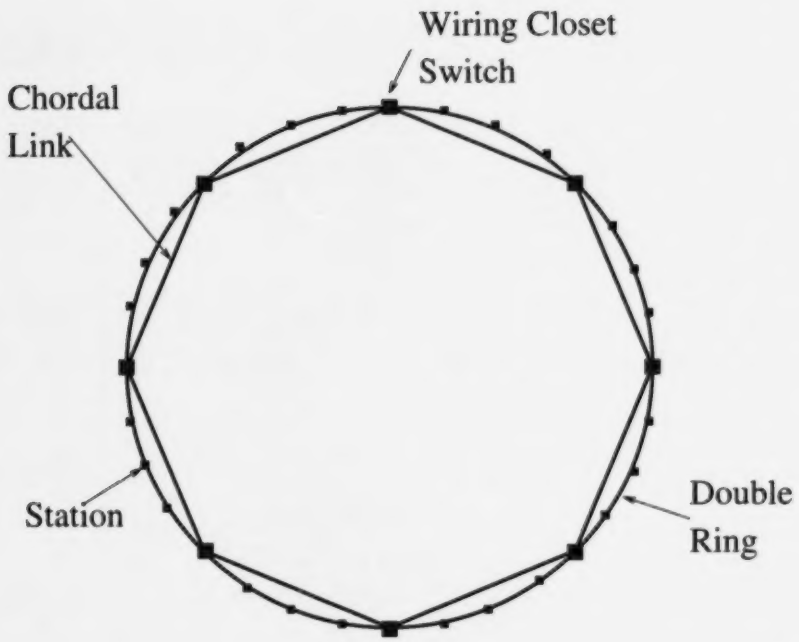


Figure 3.13: Chordal ring architecture.

Chapter 4

Mac protocols

In a *Local Area Network* (LAN) or a *Metropolitan Area Network* (MAN), all stations share a common transmission medium. Clearly, some means of controlling access to the medium are needed, otherwise devices unavoidably interfere with each other transmissions. A *Medium Access Control* (MAC) protocol determines in which way the common transmission medium is shared. A general introduction to MAC–level protocols can be found in [Sta87]; a more detailed survey is given in [RS90, Hal92]. This section describes various MAC–level protocols for fiber optic networks investigated in this thesis. Generally they are divided into two parts:

- Protocols for bus architecture.
- Protocols for ring architecture.

As mentioned before, each description refers to a *dual bus* architecture and a *dual ring* architecture. These architectures, when used with appropriate protocols (using shortest path routing), yield much better performance than other protocols lacking this feature. It is especially important for ring shaped networks where protocols like FDDI using only one ring (a second ring although present, is used for crash recovery only) cannot compete with protocols like METARING even if the transmission speed of network devices is identical. In spite of its deficiencies, FDDI is briefly discussed in Section 4.2.

4.1 Protocols for the bus architecture

This section describes some of the most important MAC–level protocols for the dual bus architecture. The SIMPLE protocol is described first, because of its simplicity, and it is combined with *reuse protocols* described in Section 4.1.3. DQDB follows as the most widely known and researched bus protocol. The section ends with the description of an interesting enhancement of the bus and ring protocols which enables a partial destination reuse of received slots. Sections 6.1.1 and 6.2.1 present performance results for some of the protocols discussed in this section.

The dual bus architecture is a very popular architecture for both LAN's and MAN's. There are several reasons for this popularity. Firstly, the physical layout of existing networks often resembles a bus with many attached stations. Secondly, bus shaped networks have no problems identifying and eliminating mis–addressed packets (slots). Unlike in ring networks where special care must be taken to trace and remove slots addressed to non-existing stations, in bus networks such faulty packets are removed automatically upon reaching the end of the bus.

For protocols which for routing purposes require information about the physical layout of the network, bus networks are easy to describe. The distance between two arbitrary stations on a bus network is easily obtained even though the exact distances between stations are not known a priori [DGR91].

The architecture of a network used by all these protocols is shown in Figure 3.7. In this configuration, all stations are connected to two unidirectional busses *Bus1* and *Bus2*. Each station is equipped with two active or passive taps connected to the fiber (Figure 2.2). All stations are numbered in a consistent way, i.e., from left to right or in the opposite direction.

The tap allows the station to read and write information to the fiber link with or without latency. Both operations can be performed at the same time. The latest property of the tap is very important, because it allows the station to test and set bits in an incoming slot on the fly.

Although unslotted operation for bus architecture is possible, slotted operation

is preferred. Unlike in the unslotted version, stations can operate without buffering incoming packets. It is also the preferred method of organizing information in contemporary networks, as specified by ATM.

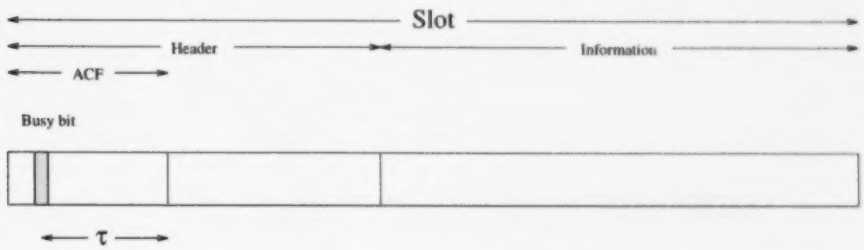


Figure 4.14: The slot structure.

Figure 4.14 describes a typical slot structure. In general a slot consists of a header and a payload part. The header usually carries several bits of information like the source and destination addresses, priority, and checksum(s). It also includes the so-called *busy bit* which is used to differentiate between empty and full slots. If a station wants to transmit a packet it must wait for a slot with the busy bit cleared. If it receives such a slot, it sets the busy-bit and inserts the packet into the payload part of the slot.

Because the station needs a certain amount of time τ to determine and possibly change the status of the incoming slot, a delay must be introduced between the busy bit and the part of the slot which should be modified if the station decides to use this slot. A careful layout of the header usually minimizes the amount of wasted space.

If a station wants to transmit packets to another station it must choose the bus with the right direction. It means that for a given station one tap is used to send packets to the left and the other to send packets in the opposite direction. As is easy to see, only for the station in the middle of the bus the average distances to the stations on the left and right are equal. It means that the bus architecture is asymmetric and potentially unfair.

Another and perhaps more unfriendly property of the bus architecture is the

starvation potential. If an upstream station has a steady stream of information to send to downstream stations, it can easily starve all downstream stations. To avoid such a mishap many ingenious techniques have been proposed.

Several access protocols have been proposed for the bus architecture. Their behavior is well known and described in many papers. The most famous of them are DQDB [DQD91], SIMPLE [Lim90], FASNET [LF82] and CRMA I [Nas90].

4.1.1 Simple

The SIMPLE [Lim90] protocol belongs to the simplest but very effective starvation prevention mechanism. It works in the regular dual bus architecture (see Figure 3.7) with two slot generators and a number of regular stations connected to two unidirectional busses.

The protocol works as follows. Each station maintains two counters denoted *S_Count(i)*, each counter is responsible for a given direction. A station can send its packet upon receiving an empty slot. Every time the station manages to send its packet, the counter corresponding to the traced direction of the packet is decremented by one. If the value of the counter reaches zero the station must ignore all empty slots traveling in this direction.

The value of *S_Count(i)* is set to *S_Priority(i)* [1] when the station receives a *Start of Cycle* signal from the opposite direction bus. This signal is generated by the slot generator when it receives an empty slot.

The event of receiving an empty slot by a slot generator can be interpreted in one of the following two ways.

- All stations willing to send their packets have reached their limits as determined by *S_Count(i)*.
- The traffic is very light and although some stations have not reached their

[1] The value of *S_Priority(i)* is determined at the initialization time of the network and can be different for each station.

sending limits, none of them have packets to send.

In both cases the slot generator must start a new cycle. The new cycle signal is usually marked by setting a special bit in the header of an empty slot inserted into the network.

Upon receiving the slot with the *Start of Cycle* bit set to 1, the station sets the value of *S_Count(i)* to *S_Priority(i)* and is immediately allowed to send its packets.

Because downstream station will receive the *Start of Cycle* bit earlier than upstream stations (the *Start of Cycle* bit propagates from downstream stations to upstream stations), SIMPLE protocol helps to evenly distribute access between all stations connected to the bus.

4.1.2 DQDB

As we can see from analyzing the SIMPLE protocol, some slots must be left empty even though some stations have packets to send. The Distributed Queue Dual Bus (DQDB) protocol [DQD91] avoids this problem by using a different (distributed) network access scheme.

The Access Control Field (ACF) in the header of the DQDB protocol consists among others of a *request bit* and two *empty bits*. Two empty bits are needed because DQDB supports regular and prearbitrated slots (for the sake of simplicity, only regular slots will be considered).

To send a packet a station must first reserve a slot on a bus where the destination station is downstream. To do so, the station examines the opposite bus, looking for a slot with cleared request bit, and sets it to 1. The request bit indicates to upstream stations that there exist a downstream station ready to send its packet. At the same time the station observes the bus where the request bit has been set, and increases a private request counter every time it sees a slot with request bit set. After the station reserves a slot by setting the request bit, it stores the value of the private request counter and resets its value to zero. The stored value describes how many empty slots

the station must ignore before it is allowed to transmit its packet. This mechanism is used in both directions so two copies of private request counters are needed.

Although DQDB appears to function better than SIMPLE it also has its problems. Extensive analyses conducted on the performance of DQDB [CGL91, DM90, Won89] have shown that some of the stations in the bus have better access to the medium than others. One of the countermeasures is *Bandwidth Balancing* (BWB) [vA90].

4.1.3 PSI

Simple Reuse Protocol using Previous Slot Information (SRP/PSI)

DQDB, SIMPLE and CRMA I cannot reuse received slots. A station cannot decide whether the slot has been received or not because it cannot buffer the header, and the busy bit alone is insufficient to make such a decision.

The introduction of an additional bit called the *reuse bit* makes it possible to reuse some of the received slots [SS93].

The structure of a new slot is depicted in Figure 4.15. Now two decisions must be made during the interpretation of the header information, so the size of the header must be increased appropriately, to allow for decoding and processing delays (τ).

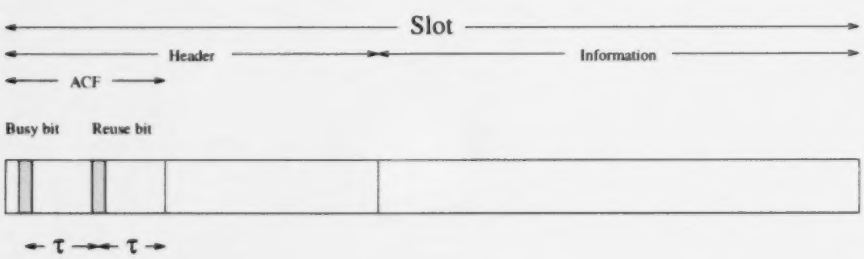


Figure 4.15: PSI slot structure.

The reuse decision is made in two stages. First the station which is about to use an empty slot checks if the destination of its packet is the same or closer than the destination of the previous slot (which just passed the station). If it is so, the

station clears the reuse bit of its own packet before it is sent. The reuse bit carries the information that the slot can be potentially reused by another station. If the destination of the previous slot is closer than the destination of the current packet the station sets the reuse bit to 1. Such a slot cannot be reused.

The second stage of the reuse process occurs at the station which tries to reuse the incoming slot. The station knows the destination of the previous slot and compares it with its own address. If the previous slot has already been received by its destination, the station assumes that the incoming slot can be potentially reused. First the station reads the busy bit. If it is cleared no reuse is necessary and the station tries to determine if the packet can be marked as reusable as described above (first stage). If the busy bit is set, the station reads the reuse bit. If the bit is set the station notices its destination and waits for the next slot. If it is cleared the station knows without looking at the destination address that this slot has been also received by its destination and it can be safely reused. The station sets the reuse bit to 1 as for the first stage and replaces the previous contents of the slot with its own packet.

Optimized Reuse Protocol using Previous Slot Information (ORP/PSI)

The Simple Reuse Protocol allows some slots to be reused. The efficiency of this protocol depends on the order in which slots are being sent. To maximize the probability of reuse some protocol modifications are proposed here.

The modifications concern mainly the way messages are handled at the station. First the station keeps not one but a number of queues with messages to be sent. Each queue keeps messages addressed to one destination. It allows the station to quickly choose the message which maximizes the reuse possibility when the station is about to send its packet, as follows.

The reuse algorithm is slightly altered to make use of the multiple queues at the station. If the station is about to reuse the incoming slot it tries to use a packet destined to the station closest to the next erasure node. This way the slot (which cannot be reused again) can be emptied by the erasure node and used again as soon as

possible. If the station cannot find any appropriate packet fulfilling this requirement it looks for the packet closest to the second erasure node. If such a packet still does not exists, the station looks for the packet whose destination is the farthest.

In case the station is about to use an empty slot it tries to find a packet destined to the same or a closer station than the previous slot. This allows the station to clear the reuse bit in the slot header and increases the probability that the slot will be reused. If such a packet does not exist the station chooses the packet which is addressed to the farthest station. This way it starts a new cycle allowing other stations to use the optimized reuse algorithm more effectively.

4.2 Protocols for ring architecture

This section describes some of the MAC–level protocols for the ring architecture. First, FDDI is described as an example of a ring protocol for medium speed networks (100 Mbps). Next METARING is described as an example of a contemporary protocol for very high speed networks (1 Gbps and up). Finally CSMA/RN Ring [FMO+91] protocol is presented as another non–slotted protocol suitable for very high speed networks.

4.2.1 FDDI

The Fiber Distributed Data Interface (FDDI) [Ros89, DB88, FDD90, FDD87, Jai89] network is defined in ISO 9314 and available as a commercial product from many vendors. It operates at the bit rate of 100 Mbps and is usually used as a back-bone network in large institutions. It uses two counter rotating rings, one called *primary ring*, the other called *secondary ring*. The secondary ring is used as a back–up when a malfunction occurs in the primary ring. Figure 3.12 shows the typical configuration for FDDI network.

FDDI uses variable length frames with the frame format shown in Figure 4.16. Apart from frames carrying information, a special information-less frame called *token*

Information frame

PA	SD	FC	DA	SA	INF	FCS	ED	FS

Token

PA	SD	FC	ED

PA - Preamble
SD - Start delimiter
FC - Frame control
DA - Dest. address
SA - Src. address
FCS - Frame check seq.
ED - End delimiter
FS - Frame status
INF - Information

Figure 4.16: FDDI frame format.

is used to synchronize transmission.

Two types of traffic are supported by FDDI: stations are able to send both *synchronous* and *asynchronous* data. The amount of synchronous data a station is allowed to send when it holds the token is determined at the network initialization time and is known as the *Synchronous Allocation Time* (SAT).

At the time of network initialization, another constant called the *Target Token Rotation Time* (TTRT) is set. This constant describes the maximum time the station has to wait to be able to send its packets. Each station is equipped with two timers—one measuring the *Token Rotation Time* (TRT) and the other, called *Token Holding Timer* (THT), measuring the time the token is held by the station.

FDDI operates as follows. Upon receiving a token the station copies the current value of the TRT timer into the THT and sets the TRT to zero. From now on the values of both timers are concurrently updated. The station also constantly computes the difference between the THT and TTRT. Next, the station sends its synchronous data trying not to exceed the SAT. If the difference of the current value of THT and TTRT is larger than zero and the station has packets to send, it is allowed to do so until

the difference between THT and TTRT reaches zero. When this happens or when the station runs out of packets to send, it releases the token by sending it to the next station.

As can be seen, no station can transmit its packets when the token travels from one station to the other and the station must wait with sending its data until it receives the token. This mechanism works well when the ring is not very long, but for long rings the time the station must wait for the token is unacceptably long. Some researchers introduced the family of *weak token protocols* [DG92a] where a station does not have to hold the token to be allowed to transmit. These protocols perform better than FDDI for long (or fast) rings.

4.2.2 Metaring

To achieve the maximum throughput in the ring architecture the METARING [CO89b, WOS92] designers have chosen a simple and well–known approach, the spatial reuse of bandwidth. It is achieved with the help of cut-through buffers in stations. The architecture details of a *meta station* are described in Chapter 2.

METARING architecture allows many stations to access the ring concurrently. This way the network achieves the maximum throughput, but it is also starvation-prone. The problem of starvation will be discussed later in this chapter.

There are two version of the METARING protocol. One of them assumes that the network is filled with equally sized slots which can be filled by stations with payload. There is a special *busy* bit in the slot header that signals whether the slot is carrying a payload. A station must wait for an empty slot to start transmitting its packet. A buffer of at least the slot header size is used to facilitate this task.

The second version of the protocol assumes that packets in the network are of different sizes and can be sent at any moment, if the station senses silence in the ring. Because METARING guarantees that once a packet is inserted into the ring it will be delivered (assuming no hardware malfunctions), every stations must be equipped with an insertion buffer needed to absorb an incoming packet when the station is

transmitting. The buffer must be at least of the maximum allowed packet size.

The medium access protocol of slotted METARING, called the *Meta* protocol in the thesis is described as follows:

- Receiving. The station waits for the beginning of a slot. When it appears, the station receives it into the buffer and checks for the busy bit. If it is set, the station checks the receiver address and determines if the packet is addressed to it. If it is, the payload is received and the busy bit in the slot header is cleared. If not, the station retransmits the slot with its header intact.

- Transmitting. The station waits for the beginning of a slot. When it appears the station starts receiving the header looking for the busy bit. If the bit is cleared, the station modifies the header with new header data. In this process the sender and receiver address is changed, and the busy bit is set.

  If the busy bit of an incoming slot is already set no action is taken and the station waits for the next incoming slot.

4.2.3 CSMA/RN Ring

METARING in its unslotted version has one major problem. To solve the packet collision problem each station must be be equipped with an insertion buffer capable of holding a maximum size packet. If the maximum packet length is large, the cost of buffering it can be very large.

To avoid large costs resulting from the need for large insertion buffers and still to be able to solve the collision problem the designers of *Carrier Sense Multiple Access / Ring Network* (CSMA/RN) protocol [FMO+91] have proposed another solution.

In this protocol (which is very similar to METARING) each station receiver is equipped with a small buffer holding about 100 bits. When a collision occurs, the station aborts its packet transmission ending the aborted packet with a truncation markers and retransmits the incoming packet. The aborted packet received at the

destination station is not discarded, but its payload is buffered and saved to be completed later when the rest of the packet arrives as a separate transmission.

This approach guarantees that the ring is not extended too much with large insertion buffers as is the case with unslotted METARING. On the other hand, the large number of aborted packets can (by introducing of a large number of header—truncation mark fields in the ring) decrease the overall throughput of the network.

4.3 Starvation Prevention Mechanisms

As mentioned earlier, if all stations are allowed to transmit at any time, as it is in the case of METARING, then a starvation can occur. Figure 4.17 depicts such a situation. In this case, station $\mathcal{A}$ sends a stream of data to stations $\mathcal{C}$ and $\mathcal{D}$. Station $\mathcal{B}$ is starved: it does not get empty slots which it could use to send its own data.

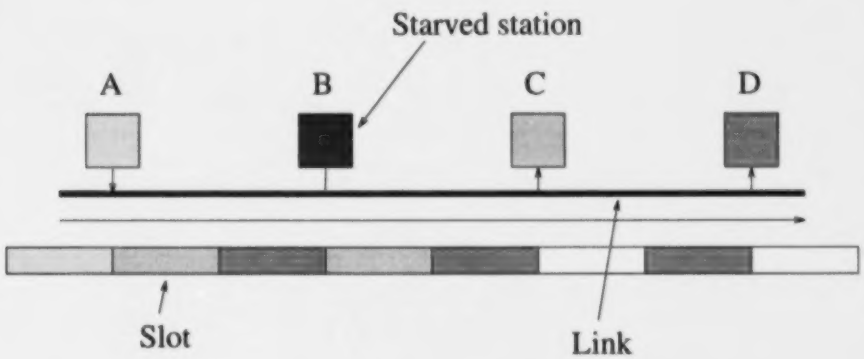


Figure 4.17: Starvation in METARING.

To avoid starvation, a station must meet additional requirements before using empty slots. Tree different mechanisms have been proposed for the METARING protocol to prevent *starvation*. All these mechanisms can be applied to other protocols as well.

4.3.1 SAT mechanism

The original METARING protocol is enhanced with a SAT mechanism [CO89a]. To prevent starvation two additional notion are introduced:

- SAT token(s)
- Slot counters

Together they provide a mechanism that works as follows (see Figure 4.18). At the beginning network operation the slot counter of every station is set to *Limit(i)* – a constant which can be different for different stations. This value limits the number of slots the station can send before it gets a SAT *token*.

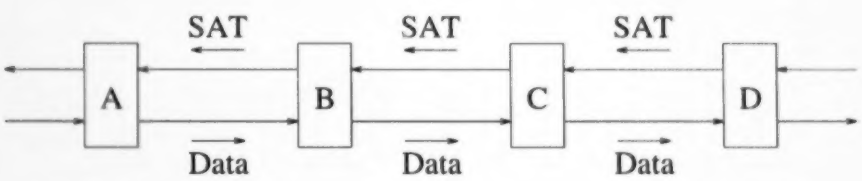


Figure 4.18: SAT–based Global Fairness Mechanism (GFM).

Each time a station wants to send a packet it examines the value of the *slot counter*. If this value is positive, the station goes ahead and puts its data into the first available empty slot. Then the counter is decremented by one. If the value of the slot counter is zero the station ignores all empty slots until it gets a SAT *token*.

The SAT token(s) are inserted into the rings during the initialization phase of the network. Each ring receives at least one SAT token. When a station receives a SAT token it sets the value of its slot counter for the given direction to *Limit(i)*. Next the station sends a SAT token to its upstream neighbor. If the value of Limit(i) is big enough this mechanism guarantees almost unrestricted access to the medium under uniform traffic conditions.

If a station feels starved (based on the count of unsent packets or when a certain time interval expires since the station was last allowed to send) next time it receives

a SAT token the station keeps it. The token is kept until the station has sent the desired[2] number of packets.

Because holding of a SAT token by one station may cause other stations to reach their limits of sent packets, this mechanism guarantees that the station will eventually get the desired part of the bandwidth.

Experiments have shown that sending SAT tokens in the direction opposite to the direction of the regulated traffic gives better results. This is because the station causing starvation is sooner affected by the absence of a SAT token.

4.3.2 Local Fairness Mechanism

Although the SAT *mechanism* works fine for uniform traffic it does not perform well for more correlated traffic patterns. The SAT mechanism is a global fairness–enforcement policy and it can penalize stations which do not contribute to starving a particular station.

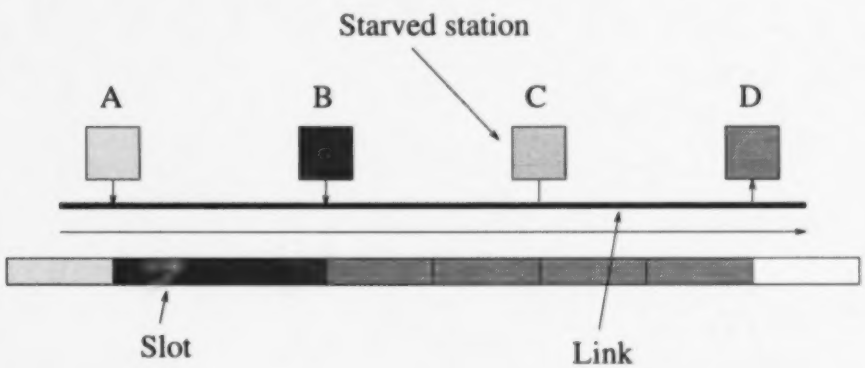


Figure 4.19: Local starvation on a bus.

Figure 4.19 presents such a case. Station $\mathcal{B}$ sends its packets to station $\mathcal{D}$ starving station $\mathcal{C}$. Station $\mathcal{A}$ which sends its packets to station $\mathcal{B}$ (and does not starve station $\mathcal{C}$) will be blocked by the SAT mechanism although its packets never reach station $\mathcal{C}$.

[2]This value is either equal to *Limit(i)* or to another value determined at the initialization time of the network.

To avoid such unnecessary side-effects, a new protocol regulating access to the medium was proposed. This mechanism is a *local fairness* mechanism and its performance should not affect stations which do not participate in the starvation of a station or a group of stations.

The station using the LFM (Local Fairness Mechanism) [CCO93] works in two modes:

- Non-restricted mode
- Restricted mode

In the first mode, any station can transmit its packets as long as it obeys the general access rules of the spatial reuse protocol, i.e., its insertion buffer is empty. If the traffic is uniform this is the state where the station supposed to spend most of the time.

In the second (restricted) mode the station can only send a preset number of packets before its state changes to unrestricted mode.

The state changes are triggered by receiving two signals:

- REQ: A signal which causes transition to restricted state.
- GNT: A signal causing transition from restricted to unrestricted state.

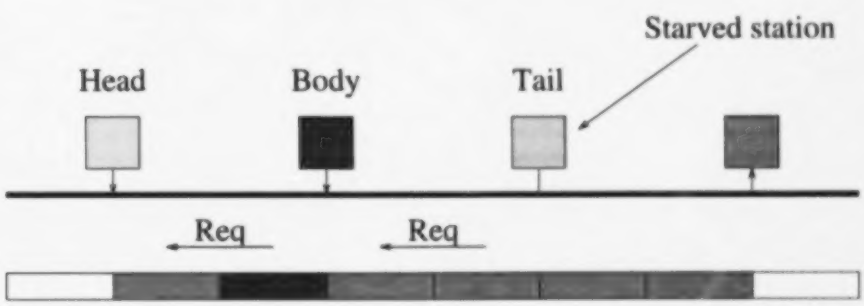


Figure 4.20: LFM on a bus.

Figure 4.20 presents a bus fragment in which a starvation occurs. This part of the bus can be treated as a local resource and the stations form a chain consisting of

the *tail*, the *head* and some *body*. The tail station is starved and the body and head stations contribute to this starvation. None of the upstream stations starting from the station next to the head station is involved in the starvation.

If the station senses that it is starved, it sends a REQ signal upstream. At this moment the station becomes the *tail* station. Upon receiving this signal, the next upstream station puts itself into restricted mode and if its upstream stations are idle (there is no traffic from them) it becomes the *head* station. Otherwise it sends a REQ signal upstream and becomes a *body* station.

At the end of this process the *body* and *head* stations have entered restricted mode of operation and from now on they are allowed to transmit only a restricted number of slots. This guarantees that after a certain well defined period of time the tail station will get a chance to transmit its packets. When this happens, the tail station sends a GNT signal upstream. After a body station receives the GNT signal it becomes the tail station. It can re-send the GNT signal to its upstream neighbor or, if it is starved, it can remain the tail station and wait for its turn to send its packets. After receiving the GNT signal, both body stations and the head station transit to unrestricted mode. The propagation of the GNT signal stops at the head station.

Figure 4.21 presents a state transition diagram for the LFM mechanism.

The mechanism described above presents a simplified version of the LFM protocol. It assumes that only one starved station exists and it initiates the transition from unrestricted to restricted mode of operation in a part of the network. It is not the only possible starvation scenario.

If there is more than one initiator of the starvation prevention procedure, i.e., two or more stations sense starvation and send REQ signals, two variations of the algorithm are possible:

- To sequence REQ requests
- To merge REQ requests.

The first possibility is less effective because it can lead to a quadratic time-out

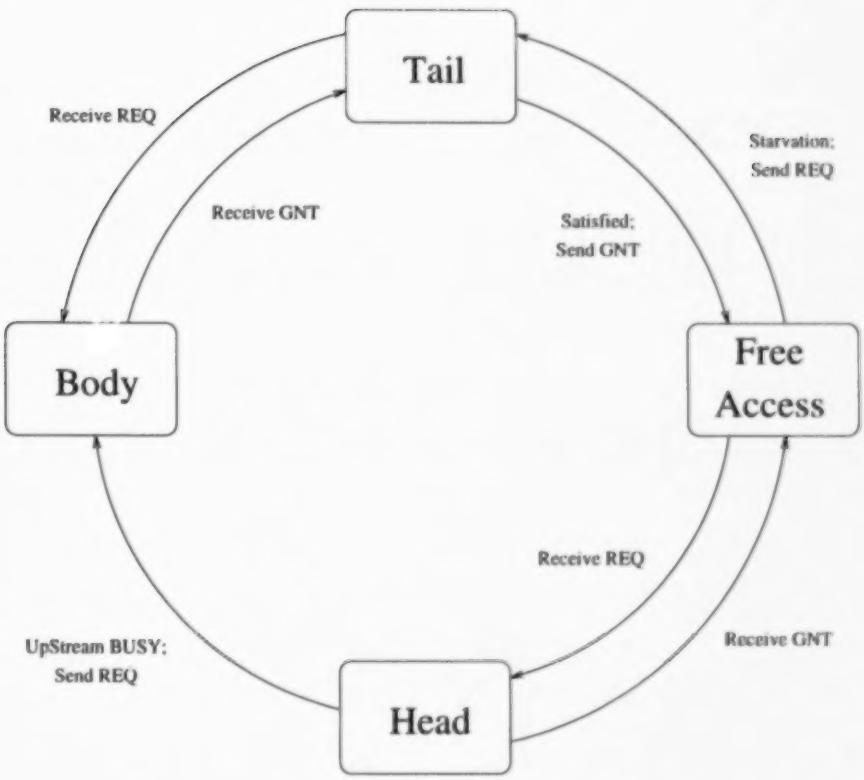


Figure 4.21: LFM state transition diagram (bus architecture).

bound [CO89a]. On the other hand, if the second modification is chosen the tail station must be prepared to change its status from tail to body upon receiving a REQ signal. Figure 4.21 takes into account the latest enhancement of the algorithm.

If the second option is used, one more problem shows up. In the ring architecture it is possible that the request path will spawn the whole ring and all stations would transit to *body* state. If this is not recognized, it will lead to a deadlock.

As long as the *request path* consist of one *tail* and *head* station, deadlock is impossible. To enforce that, the REQ signal is augmented with a field that carries the identification of the tail station. Additionally, each station is equipped with a variable REQ_ID which holds the number of the tail node that originated the starvation prevention procedure.

A tail station receiving a REQ signal must compare its REQ_ID with the id number carried by the signal. If these values are not equal, the two *Request paths* are merged and the station state changes from tail to body. The station's REQ_ID must also be updated and the new REQ re-send if the REQ_ID is smaller then the REQ id.

4.3.3 Modified Local Fairness Mechanism

Although *Local Fairness Mechanism* is much more versatile than the SAT mechanism, it has its own problems. Consider the situation presented in Figure 4.22. Stations $\mathcal{A}$ and $\mathcal{C}$ send their packets to station $\mathcal{E}$ starving station $\mathcal{D}$. Although station $\mathcal{B}$, sending its packet to station $\mathcal{C}$, does not participate in starving station $\mathcal{D}$, it will nevertheless transit to restricted mode of operation and become a body station.

This happens because even though the *tail* ID is included in the REQ signal, it is not used to determine whether the station that receives it participates in the starvation process. A simple change in the starvation prevention algorithm fixes this problem.

The REQ signal is augmented with an additional field SUF (for suffice), which declares how many packets the *tail* station needs to transit from starved to non-starved state. Upon receiving the REQ signal, the stations checks if the packets it

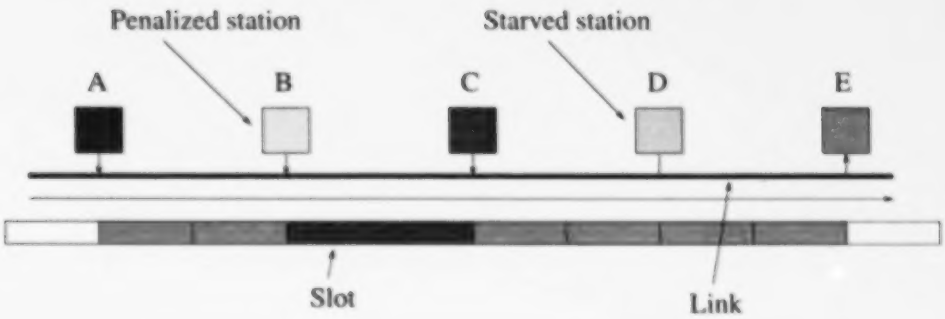


Figure 4.22: Local starvation on a bus. General case.

currently sends are addressed further away than the starved station (by checking the REQ_ID field). If this is the case, the station enters the restricted mode of operation and behaves as a body station. Otherwise, it still checks if the REQ signal should be retransmitted, but it does not enter restricted mode of operation. This way Station $\mathcal{B}$ in the example continues sending packets to station $\mathcal{C}$. In a situation where the station sends some of its packets behind the starved station, it enters a new *limited* mode of operation. In this mode the station is allowed to send only packets which do not starve the *tail* station.

At the same time, when the decision concerning the transition from unrestricted to restricted mode is made, the station loads the value of the SUF field to a countdown counter and starts counting slots passing by in the direction of the starved station. If this counter reaches zero, the station transits from restricted to unrestricted mode.

The SUF mechanism eliminates the need for the GNT signal. It also has other advantages over the non–modified LFM. If every station knows the distances to other stations in the network, additional improvements of LFM can be made. Since a station knows the id of the tail station and the distance to it, it can calculate how many empty slots the starved station has received since the REQ signal was originally sent. This way (especially if the slots are small, as in ATM) the station can automatically decrease the requested SUF value in the REQ signal. This minimizes the time the station must spend in the restricted mode and, if the distance from the tail station is sufficiently big, it can eliminate the explicit transition from unrestricted to restricted

mode completely.

Chapter 5

Simulation Environment

A series of simulation experiments were performed to investigate the property of the chordal ring and the chordal bus architectures under the uniform traffic pattern. Although analytic models are useful and tractable when uncomplicated assumptions are made, there are too complicated for analyzing real-life models. Simulation of unsimplified models provides results which can be used to test hypotheses under nontrivial conditions [Gbu93].

This chapter describes the simulation environment used in this thesis. Chapter 6 presents the results of the experiments.

5.1 The Simulator

The simulators for different network architectures were written in *C++* using SMURPH simulation environment [Gbu93]. SMURPH provides a simulator template as a simulator core and a number of classes defining real and virtual network objects.

In this environment, a network consists of interconnected stations. Each station (which represents a real network device) has a number of active processes whose responsibility is to model the operation of receiving and transmitting packets. Each process is a *finite state machine* and mimics a fragment of the MAC protocol.

The following assumptions are made in the simulation environment:

- Each message is sent to exactly one station, i.e., there are no broadcast and multi–cast messages.
- Packets are inserted into empty slots and removed by destination stations (in META protocols) or by erasure (wiring closet) nodes (in reuse protocols).
- Every slot consists of 8192 bits of payload and 100 bits of header. Every slot is delayed for at least 100 bits of time at each station to allow for the header contents examination.
- Two adjacent slots are separated by an inter–slot gap of 1 bit.
- The default bus and ring length is 100,000 bits unless stated otherwise (12 slots).
- The network consists of 36 stations evenly distributed along the channel.

Apart from stations with their active processes the simulator model ports and links. Ports are SMURPH objects which represent real life ports and are responsible for receiving and transmitting packets. Processes use ports to perform send and receive operations. Links are also SMURPH objects which simulate different communication media. In the simulators built to conduct our experiments, links simulated unidirectional fiber optic connections. The topology of the network is set by connecting links and ports.

Another SMURPH object—the packet—is used to simulate information exchange in the system. Besides the information (payload) part each packet stores information about the source and destination address, creation time, message number, access bits and some debug information, irrelevant from the point of view of the simulated protocol.

5.2 Performance Measurements

Simulations take a certain amount of time to come to the state when the model operates in a steady manner. Start–up conditions in the network are quite different

than the final steady ones. It is especially true when modeling heavy traffic conditions, because the network and the message queues at the stations must first fill up with messages to represent accurately the state of the system. The behavior of the simulated system under these start–up conditions is different and should be ignored.

To reduce the impact of the start–up conditions in the simulated system, performance measures must be taken after a *warm-up* time. After some experiments, sending of 30,000 messages was chosen at the moment when the network has reached a steady state.

5.3 Traffic Conditions

Traffic in computer networks can be classified into *asynchronous* and *synchronous* or *non-real-time* and *real-time*. In the experiments conducted in this thesis only the asynchronous type of traffic is investigated.

Asynchronous traffic is characterized by:

- Irregular arrival times to the network.
- Non-uniform message sizes.
- Non-critical delivery time.
- Sensitivity to packet loss.

File transfers, electronic mail, facsimile, terminal key strokes, client—server communication are examples of asynchronous traffic.

Uniform traffic pattern, a traffic where the message inter-arrival time is uniformly distributed, has been chosen to simulate asynchronous traffic. The traffic load was uniformly and randomly distributed in the system, i.e., each station had the same probability to send a message. Since the maximum throughput has been the main concern in the experiments performed in this thesis, the message length is constant and equal to the information length of the slot.

Chapter 6

Simulation results

This chapter presents simulation results for two different architectures:

- Plain architecture
- Chordal architecture

The results are presented in two following sections. Each section is divided into three parts presenting results for following protocols:

- Simple Reuse Protocol using Previous Slot Information
- Optimized Reuse Protocol using Previous Slot Information
- *Meta* protocol

All results are presented as *Message Access Time* versus *Throughput* graphs. To better describe the properties of protocols both *mean* and *maximum* message access times are plotted as functions of the throughput. The Message Access Time is expressed in *slots* as a measure of time. The Maximum Message Access Time is measured for the last 10,000 messages sent during the simulation.

6.1 Plain architectures

The properties of *plain architecture* are described in Section 3.2.1. In the case of *Reuse Protocols* for bus architecture three erasure nodes are used. The placement of the erasure nodes has been determined experimentally where the criterion for the optimal placement was to minimize the *Mean Message Access Time*. When the differences in the Mean Message Access Time were not clear enough, the architecture with the maximum throughput was chosen.

6.1.1 Bus Architecture

This section presents the results for the bus architecture. In all cases the dual bus architecture is simulated with following parameters:

- Bus Length 100,000 bits
- 36 uniformly spaced stations
- 3 erasure nodes
- Packet Length 8248 bits
- Header Length 100 bits
- Inter-Slot Gap 1 bit
- Uniform Traffic Pattern
- SIMPLE message limit 40

Bus with Simple Reuse Protocol (Rbus)

Figure 6.23 presents the simulation results for a dual bus architecture with a simple reuse mechanism. The placement of erasure nodes is presented in Figure 6.24 and was determined by running several experiments with different locations of erasure nodes.

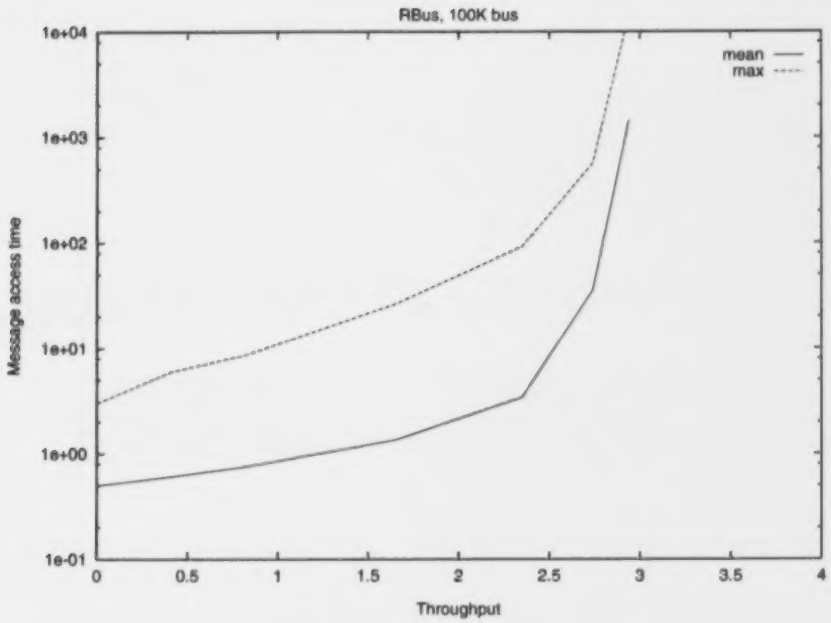


Figure 6.23: RBus performance.

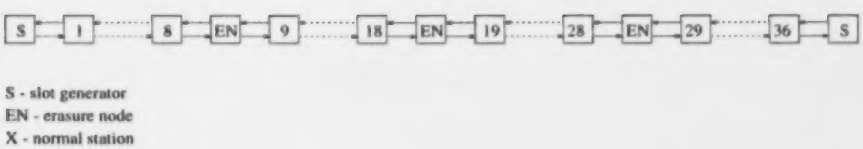


Figure 6.24: Placement of Erasure Nodes.

It is reasonable to assume that it is possible to make throughput reach the value of 3.5 by oversaturating the network, $\mathcal{T}_{real_max}$ throughput of the network is not greater than 3.0.

Bus with Optimized Reuse Protocol (ROptbus)

Figure 6.25 presents the simulation results for the same architecture as in the previous case but with ORP/PSI protocol described in Section 4.1.3.

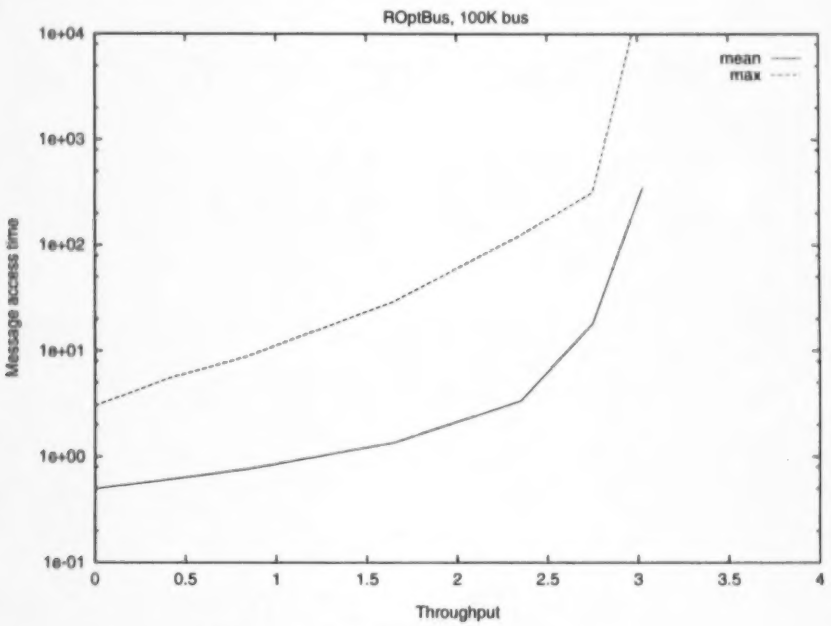


Figure 6.25: ROptBus performance.

In this case, $\mathcal{T}_{real_max}$ throughput of the network is also about 3.0. Note that the mean message access time at the throughput of 3.0 is considerably smaller than in the case of the simple RBUS architecture.

Notice that the performance of the two protocols differs only for a throughput larger than 2.2. This is caused by the fact that to change the order of the sent messages a station must have a certain number of messages in its message queue. It only happens for a relatively intense traffic in the network and the station cannot take advantage of this possibility if the load is light (the throughput is smaller than

2.2).

Bus with Meta Protocol (Metabus)

Figure 6.26 presents the simulation results for a dual bus architecture with a destination release protocol. The architecture of a station capable of destination release is presented in Section 2.2.2. Because of the nature of the META protocol no erasure stations are needed in this architecture.

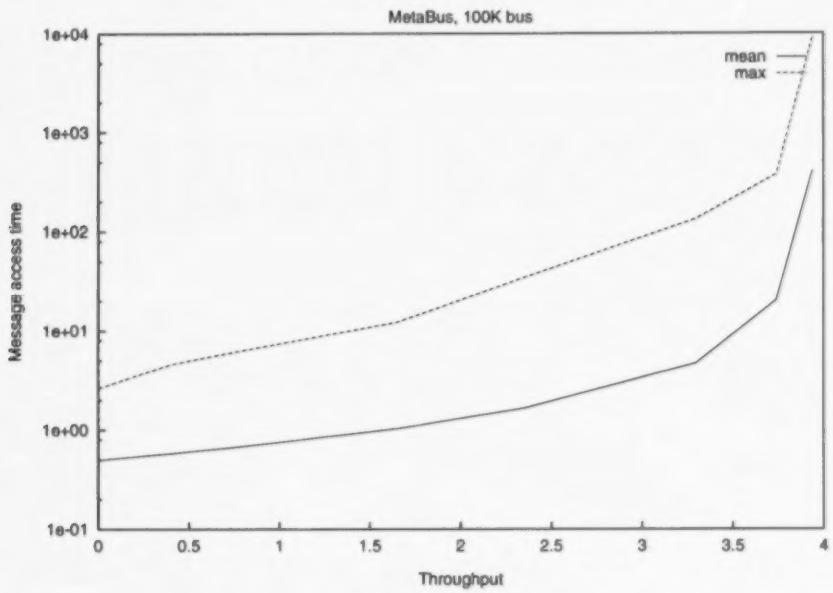


Figure 6.26: MetaBus performance.

The maximum throughput of this configuration is about 3.9. Both mean and maximum message access times are well below the 10,000 slot mark used to determine the saturation point of the network.

Notice that even for very heavy traffic (about 3.7), the mean message access time is about 10 slots which is an unrivalled result. Two previous protocols achieve similar mean message access times for throughput of 2.5 and 2.7, respectively.

6.1.2 Ring Architecture

This section presents the results for the ring architecture. In all cases the dual ring architecture is simulated with following parameters:

- Ring Length 100,000 bits
- 36 uniformly spaced stations
- 6 Erasure nodes
- Packet Length 8248 bits
- Header Length 100 bits
- Inter-Slot Gap 1 bit
- Uniform Traffic Pattern
- SIMPLE message limit 40

Ring with Simple Reuse Protocol (Rring)

Figure 6.27 presents the simulation results for the dual ring architecture with a simple reuse protocol. Six erasure nodes are placed evenly in the ring (see Figure 3.13).

The maximum throughput of the presented architecture is about 4.8. Because of the symmetry of the ring the maximum message access time does not play such important role (as in the case of the bus architecture) in determining the maximum throughput of the network. As can be seen, the shape of the mean and maximum curves in Figure 6.27 are similar and the saturation point appears for the maximum network throughput.

Ring with Optimized Reuse Protocol (ROptring)

Figure 6.28 presents the simulation results for the same architecture but with the protocol improvements described in Section 4.1.3.

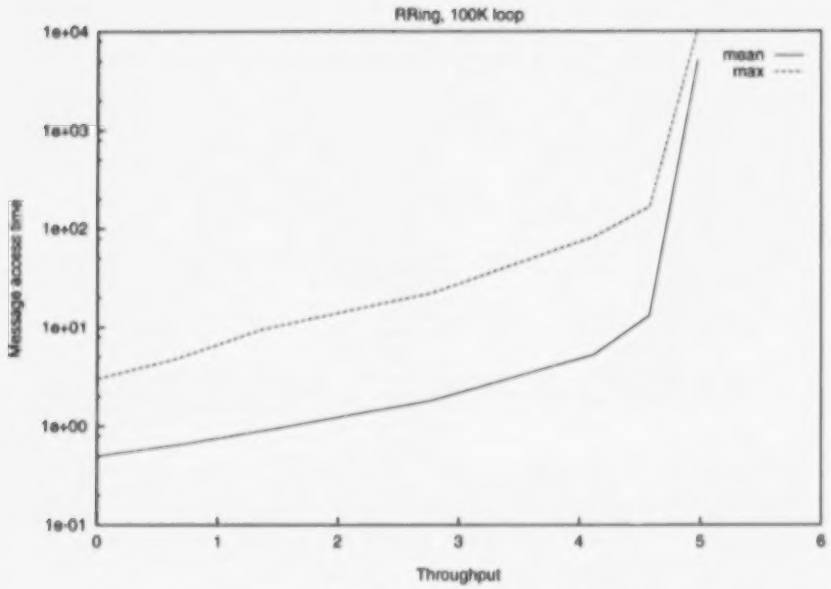


Figure 6.27: RRing performance.

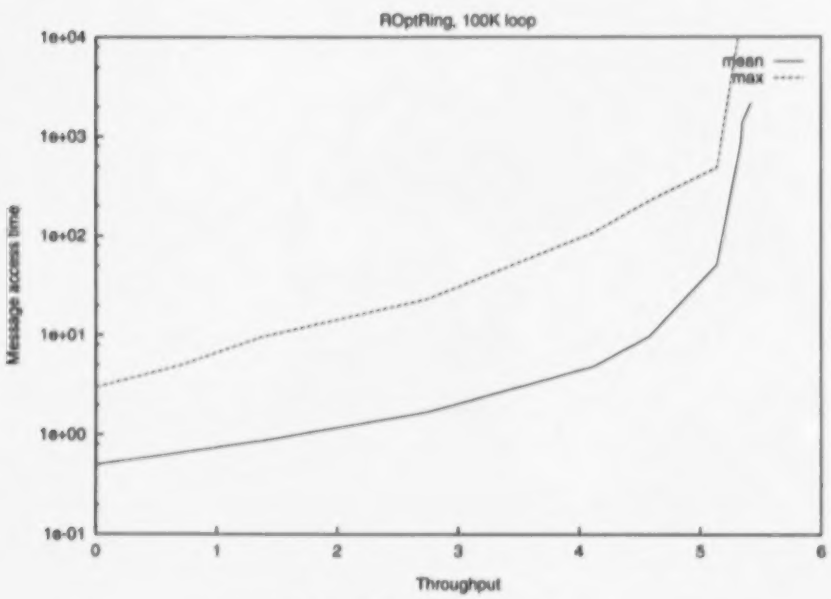


Figure 6.28: ROptRing performance.

The maximum throughput for this protocol is about 5.5 which is better than 4.8 achieved by the previous plain reuse protocol. The performance curves of both protocols are almost identical up to throughput of 4.5, then the protocol with improvements performs better. As it was described for the bus architecture, a station can efficiently reorder its message queue only if the number of messages in the queue is above a certain threshold (the traffic intensity is larger than 4.5).

Ring with Meta Protocol (Metaring)

Figure 6.29 presents the simulation results for the dual ring architecture with the METARING protocol—as described in Section 4.2.2. As with the METABUS architecture, no erasure nodes are used.

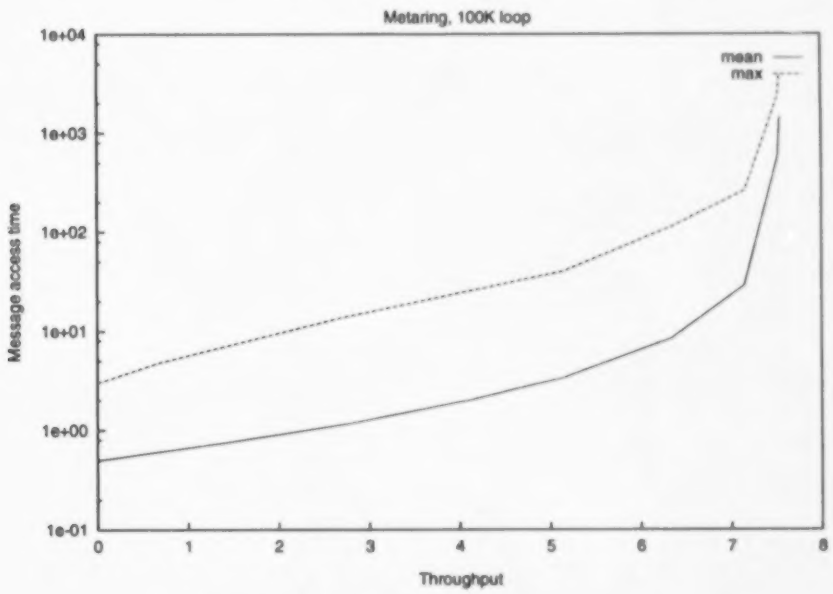


Figure 6.29: METARING performance.

The maximum throughput achieved by the METARING protocol is about 7.5. For the traffic intensity yielding the throughput of 7.0, the Mean Message Access Time is about 10 slots. This is much better than in the cases of the two previous protocols, which offer the same access time for the throughput of 4.4 and 4.6, respectively.

6.2 Chordal architectures

The properties of the *chordal architecture* are described in Chapter 3.2.2. In the case of the *Reuse Protocols* for the bus architecture (as well as for the plain bus architecture) three erasure nodes are used. Because of the additional functions played by the nodes in the chordal architecture, these nodes are called *wiring switch nodes*. The placement of these nodes has been determined experimentally with the same criteria as for the plain bus architecture (Section 6.1.1). The final placement of the wiring switch nodes is shown in Figure 6.30.

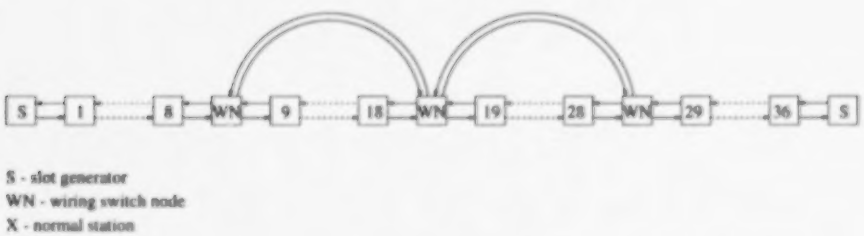


Figure 6.30: Placement of Wiring Switch Nodes in Chordal Bus.

6.2.1 Bus Architecture

This section presents the results for the bus architecture. In all cases, the dual bus architecture is simulated with following parameters:

- Bus Length 100,000 bits
- 36 uniformly spaced stations
- 3 wiring switch nodes
- Inter switch distance 4124 bits
- Packet Length 8248 bits

- Header Length 100 bits
- Inter-Slot Gap 1 bit
- Uniform Traffic Pattern
- SIMPLE message limit 40

Chordal Bus with Simple Reuse Protocol (RSbus)

Figure 6.31 presents the simulation results for the dual bus architecture with the connections between wiring switch nodes as described in Section 3.1.2.

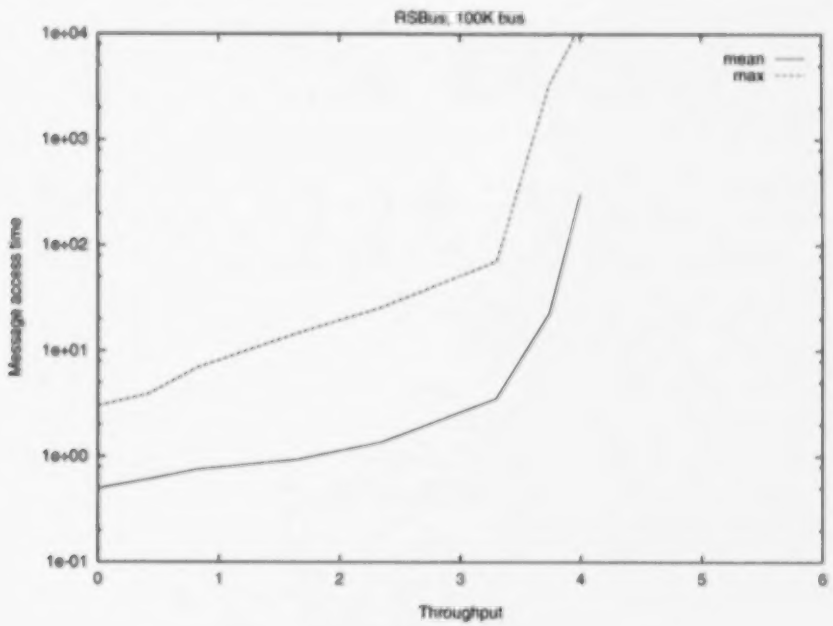


Figure 6.31: RSBus performance.

$\mathcal{T}_{real_max}$ throughput of this network is about 3.9. The maximum achievable throughput for the oversaturated network is 4.5. Notice that although the network yields the throughput of 3.9, for a slightly smaller load (throughput about 3.5) the mean message access time is still very low and equal to about 10 slots.

Chordal Bus with Optimized Reuse Protocol (RSOptbus)

Figure 6.32 presents the simulation results for the same architecture but for the optimized access protocol.

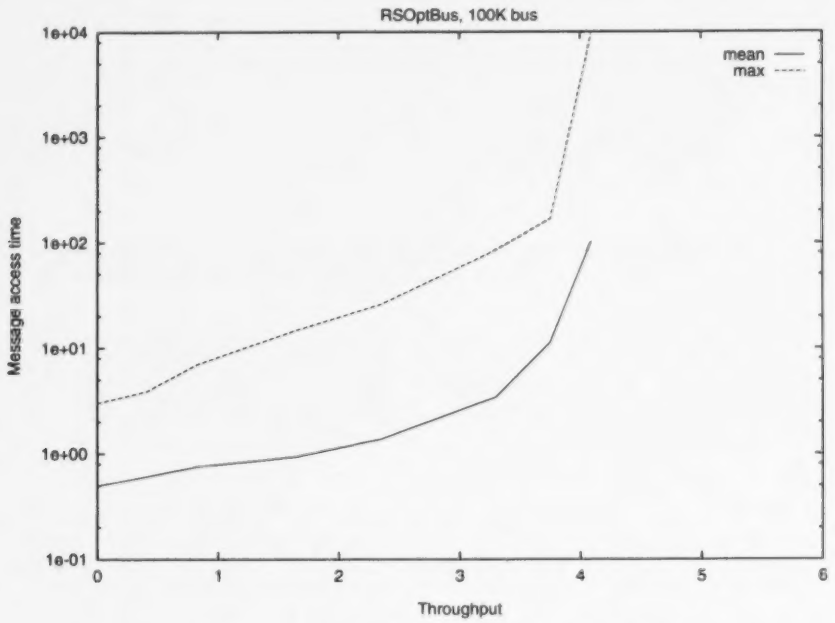


Figure 6.32: RSOptBus performance.

In this case, $\mathcal{T}_{real_max}$ throughput with maximum message access time of 10,000 slots reaches 4.0. Although it is a small increase comparing to the non—optimized version of the protocol, the throughput of 3.9 is reached with the mean message access time of 10 slots.

Chordal Bus with Meta Protocol (MetaSbus)

Figure 6.33 presents the simulation results for the chordal dual bus architecture with the META protocol. For this architecture, the function of a wiring switch node is simplified to relaying messages only.

With the META protocol, $\mathcal{T}_{real_max}$ throughput increases to 4.5. It can be seen that for the throughput of 4.3 the mean message access time reaches 10 slots.

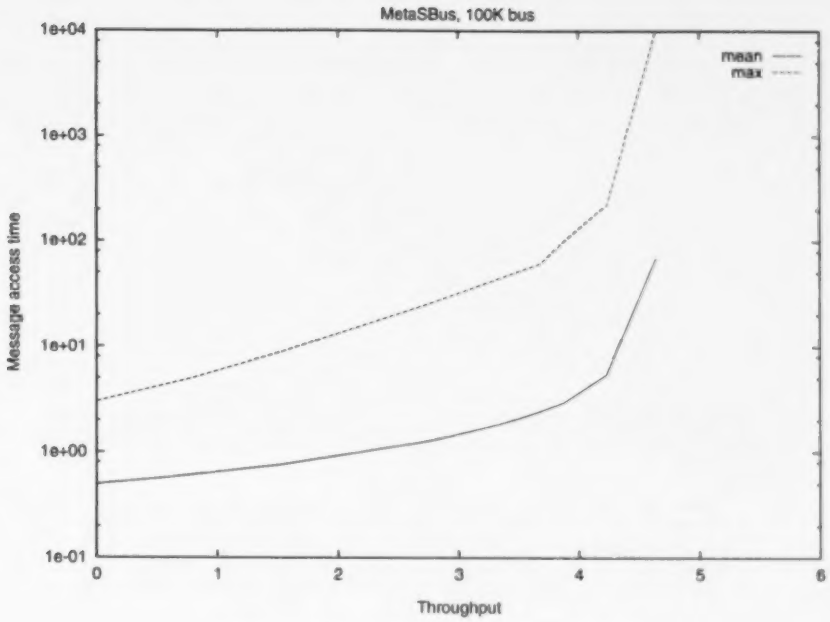


Figure 6.33: MetaSBus performance.

6.2.2 Ring Architecture

This section presents the simulation results for the chordal ring architecture. In all cases, the dual ring architecture is simulated with following parameters:

- Ring Length 100,000 bits
- 36 uniformly spaced stations
- 6 wiring switch nodes
- Inter–switch distance 4124 bits
- Packet Length 8248 bits
- Header Length 100 bits
- Inter-Slot Gap 1 bit
- Uniform Traffic Pattern
- SIMPLE message limit 40

Chordal Ring with Simple Reuse Protocol (RSring)

Figure 6.34 presents the simulation results for the double ring architecture with chordal links. In this case a simple reuse protocol is used.

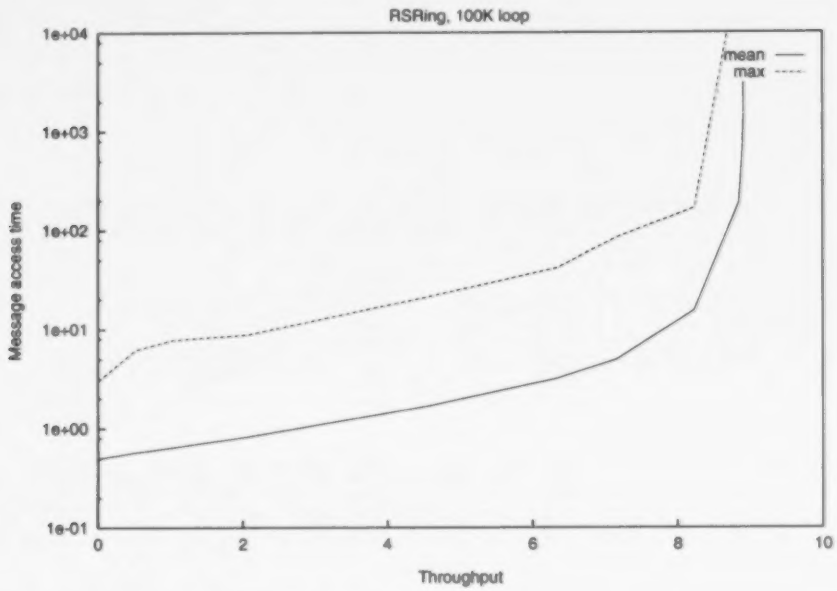


Figure 6.34: RSRing performance.

In this architecture, $\mathcal{T}_{real_max}$ throughput reached for saturated network is 8.3. The maximum throughput achieved by oversaturating the network is not much bigger and amounts to 8.4. The throughput of 8.0 is achieved with a mean message delay of 10 slots.

Chordal Ring with Optimized Reuse Protocol (RSOptring)

Figure 6.35 presents the simulation results for the same architecture as in the previous case but the protocol reorders the incoming messages in the station message queue.

As can bee seen, $\mathcal{T}_{real_max}$ throughput for the saturated network reaches 9.0. The maximum achievable throughput is about 10.0 in artificial, oversaturated conditions. A throughput of 8.2 is achieved with the mean message access time of 10 slots.

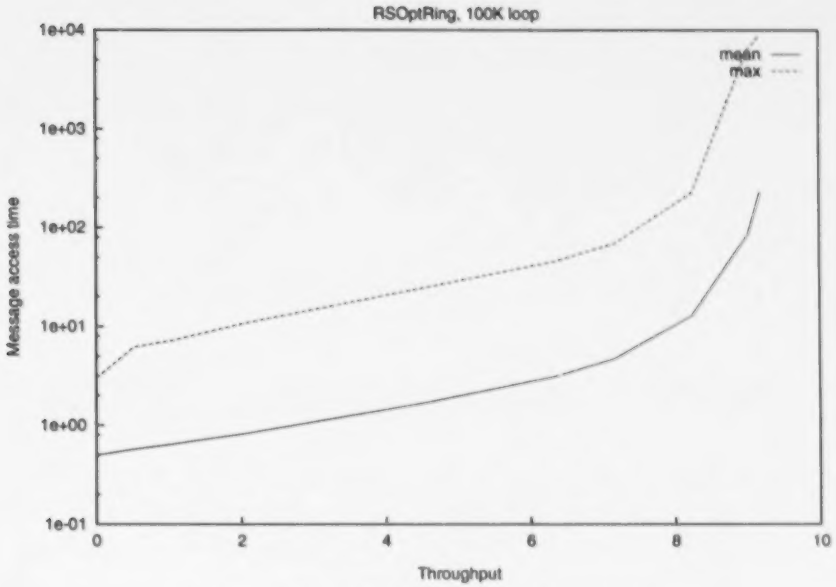


Figure 6.35: RSOptRing performance.

Chordal Ring with Meta Protocol

Figure 6.36 presents the simulation results for dual ring architecture with chordal links operating under the META protocol. In this case (similarly to the bus architecture) the behavior of the wiring switch nodes is simplified. These nodes no longer have to remove received packets from the network, and their role is limited to routing incoming packets.

In this architecture, $\mathcal{T}_{real_max}$ throughput is 13.5. As can be seen this value is reached when the network is saturated and the protocol reaches its peak performance. The protocol yields 12.3 throughput for the mean message access time of 10 slots.

6.3 Comparison

The following table presents a comparison of simulation results for plain and chordal architectures. The table presents aggregate network throughput for three different levels of Mean Message Access Time (MMAT):

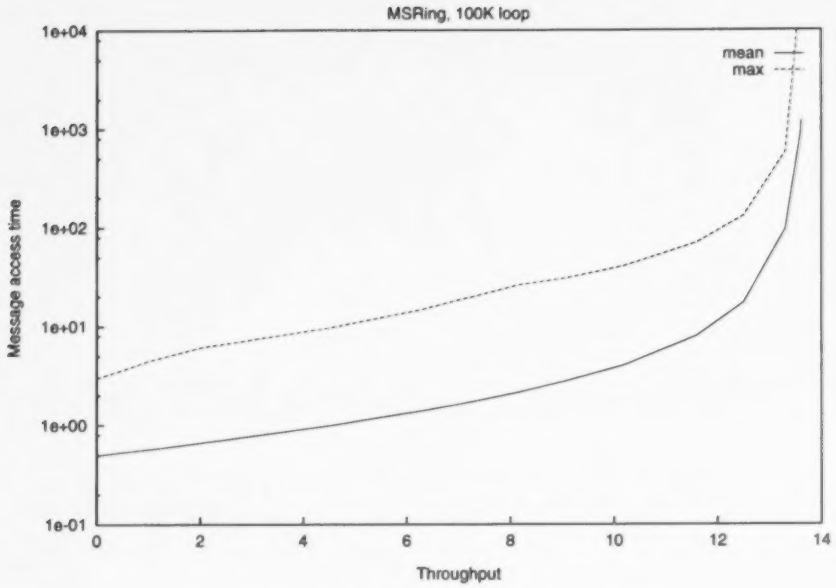


Figure 6.36: S METARING performance.

- 10—required by time critical applications.
- 10^4—realistic upper bound (maximum sustained realistic throughput).
- ∞—absolute bound.

Notice that the *Meta* protocol can be viewed as a particular (and very expensive from the hardware costs point of view) case of ORP/PSI protocol where all stations are *Meta* stations. The fractional component of the protocol name represents the relative frequency of erasure nodes or wiring closet nodes (i.e. for 36 stations, $1/12$ implies 3 erasure nodes).

The following figures present a comparison of the performance of the presented protocols in a graphic form.

Plain Architecture	*Dual Bus*			*Dual Ring*		
Protocol *MMAT:*	10	10^4	∞	10	10^4	∞
SRP/PSI— $1/12$	2.5	3.0	3.5	4.4	4.8	4.9
SRP/PSI— $1/4$	2.9	3.3	3.9	4.9	5.5	5.6
ORP/PSI— $1/12$	2.7	3.0	3.7	4.6	5.2	5.5
ORP/PSI— $1/4$	2.9	3.4	3.9	5.0	5.7	5.8
ORP/PSI— $1/1$ (META)	3.7	3.9	4.0	6.7	7.5	7.9
Chordal Architecture	*Dual Bus*			*Dual Ring*		
Protocol *MMAT:*	10	10^4	∞	10	10^4	∞
SRP/PSI— $1/12$	3.5	3.9	4.5	8.0	8.3	8.4
ORP/PSI— $1/12$	3.9	4.0	5.1	8.2	9.0	10.0
ORP/PSI— $1/1$ (META)	4.3	4.5	6.0	12.3	13.5	13.6

Table 6.1:

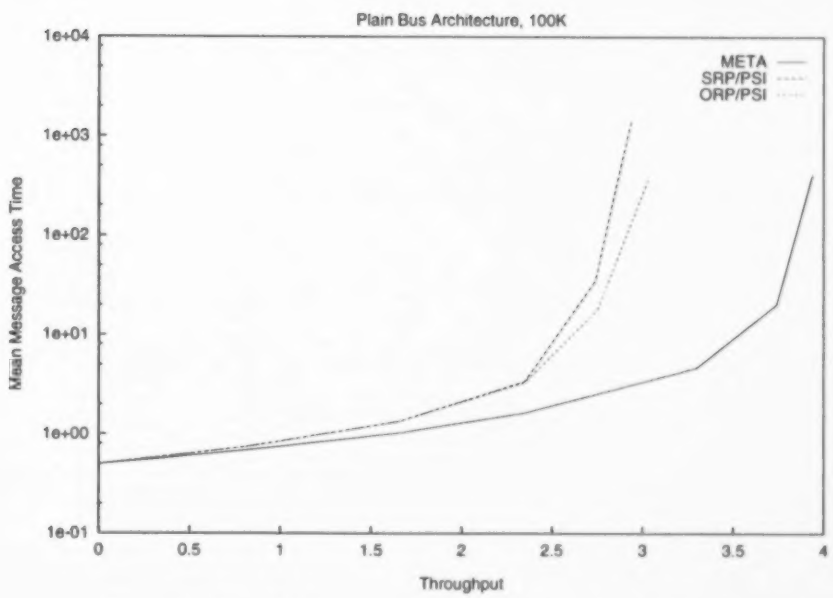


Figure 6.37: Performance of Plain Bus Architecture.

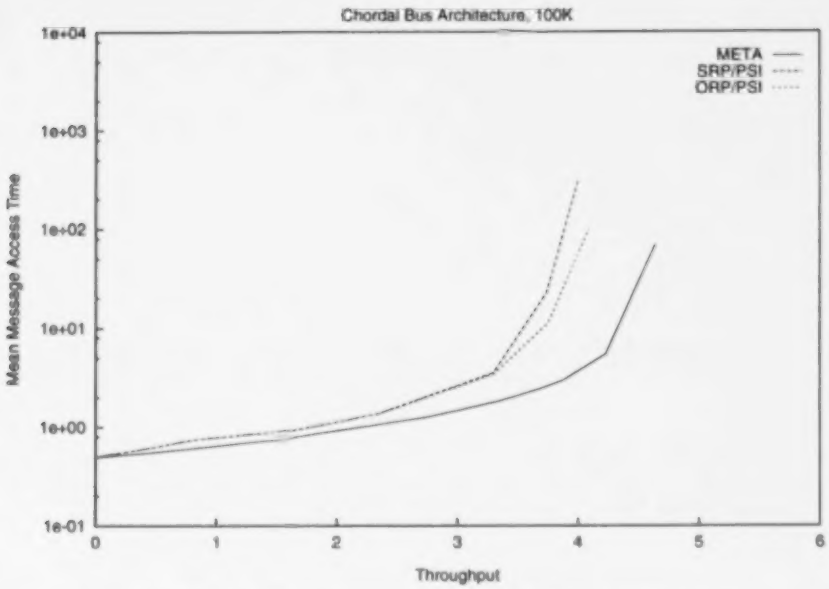


Figure 6.38: Performance of Chordal Bus Architecture.

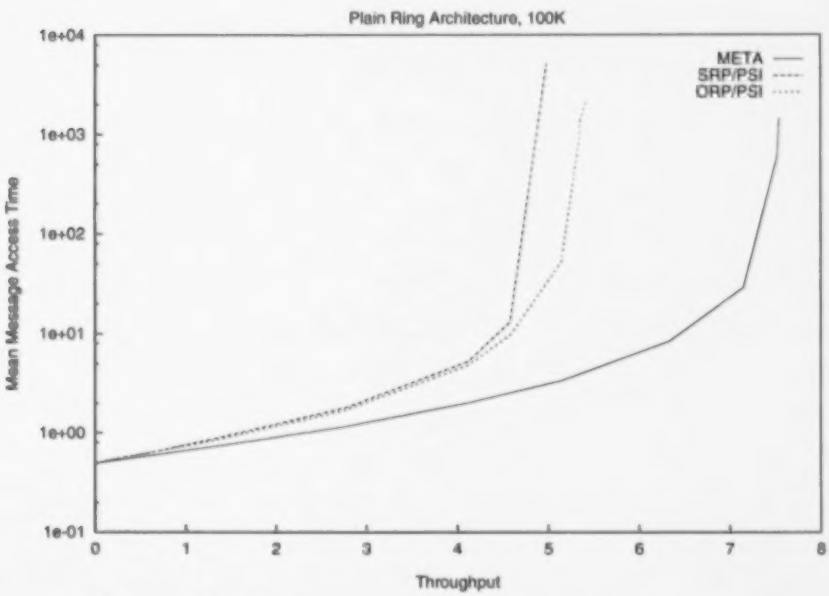


Figure 6.39: Performance of Plain Ring Architecture.

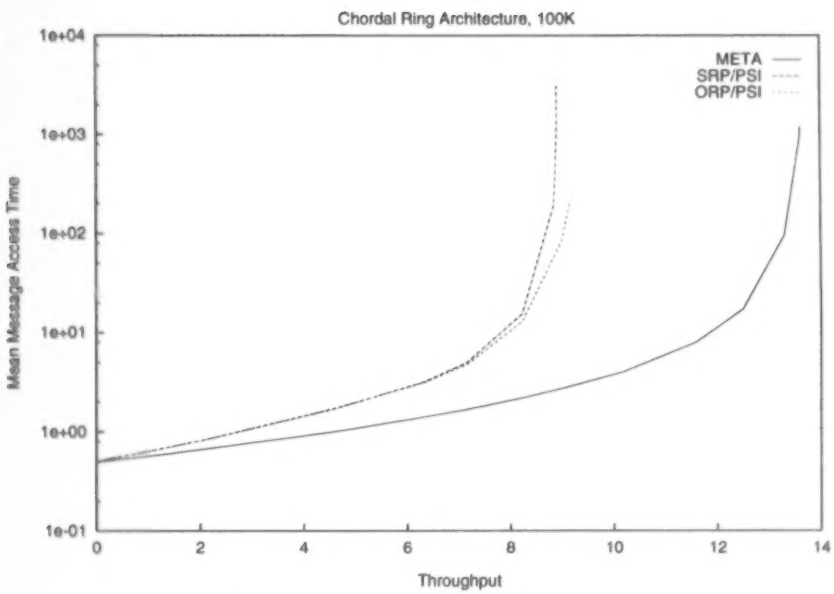


Figure 6.40: Performance of Chordal Ring Architecture.

Chapter 7

Conclusions and Future Work

The chordal bus and chordal ring topologies were described and analyzed by simulation. The properties of three Medium Access Protocols: Simple Reuse Protocol using Previous Slot Information (SRP/PSI), Optimized Reuse Protocol using Previous Slot Information (ORP/PSI) and META protocol were discussed for both the plain and chordal architectures. Simulation models were developed and experiments were conducted to investigate network behavior under uniform traffic conditions. Section 6.1 presents a detailed performance comparison of these protocols for the plain architecture. Section 6.2 presents a detailed performance comparison of SRP/PSI, ORP/PSI and META protocols for the chordal architecture. While META protocol displays the best performance of all the protocols presented in this thesis, it also requires the most expensive hardware.

ORP/PSI is an improved protocol based on SRP/PSI [SS93] and can be viewed as a generalized META protocol. ORP/PSI protocol possesses following properties:

- ORP/PSI is a *capacity-1* protocol; its maximum throughput does not depend on the network size and the transmission rate.
- The transmission rules are simple and can be implemented in hardware.
- Partial, software-supported destination release causes increased throughput comparing to traditional bus and ring protocols.

- Under light–load conditions the Mean Message Access Time is the lowest possible—half a slot.
- Small Mean Message Access Time for the wide range of network loads.
- A favorable cost—performance tradeoff when compared with META protocol, since the ratio of expensive nodes is reduced by a larger factor than the performance drop.

To prove the last point let us compare the cost of the METABUS architecture of 36 *Meta* stations yielding $\mathcal{T}_{real_max} = 3.9$ with the cost of the ORP/PSI bus with 3 Erasure Nodes and 36 regular stations yielding $\mathcal{T}_{real_max} = 3.0$. Let us assume that the *Meta* station is five times more expensive than the regular station and it costs \$5000.

The cost of METABUS system is $36 \times \$5000 = \$180{,}000$. The cost of ORP/PSI bus system is $36 \times \$1000 + 3 \times \$5000 = \$51{,}000$. Thus, for the 30% of the cost we can purchase the ORP/PSI bus system offering 75% performance of the more expensive METABUS system.

The Modified Local Fairness Mechanism (MLFM) was proposed. It possesses many characteristics of an ideal local fairness algorithm:

- MLFM dynamically divides the transmission medium into local communities of interest.
- It takes into account more realistic occurrences of starvation than Local Fairness mechanism proposed in [CCO93].
- The introduction of additional information in the REQ signal in the form of SUF field eliminates the need of GNT signal. This mechanism allows earlier transition from *restricted* to *non-restricted* mode of operation of all affected stations.

Simulation shows (see Section 6.2.2) that the chordal ring architecture is able to deliver the throughput of 13.5 when used with the *Meta* protocol[1]. It is interesting to notice that the 75% increase in the maximum throughput (compared to a double ring architecture) has been achieved with practically the same length of fiber links.

Unfortunately, such an increase in the maximum throughput is only possible when the wiring closets switches are placed evenly in the network. For a less even placement, usually encountered in real-life networks, the increase in the throughput will be less pronounced.

In this study uniform traffic pattern was used to investigate properties of the bus and ring chordal architectures. The future work should include realistic (non-uniform) traffic patterns. Beside the change in the traffic pattern model future research could concentrate on following areas:

- Large message transport—Medical Images, Multi Media Databases.
- The impact of higher layer protocols (TCP/IP, OSI) on the performance of ORP/PSI.
- Fairness in the chordal bus architecture.
- Further modifications of the chordal architecture.
- "Smart" message relaying–modified role of slot generating nodes.
- MLFM as a fairness mechanism for METARING and METABUS architectures.

[1]The Meta protocol is described in Section 4.2.2.

Bibliography

[AL81] B.W. Arden and H. Lee. Analysis of chordal ring network. *IEEE Transactions on Computers*, 30(4):291–295, April 1981.

[BDG93] C. Baransel, W. Dobosiewicz, and P. Gburzynski. CBRMA++/SR: On the design of a MAN/WAN MAC protocol for high–speed networks. *IEEE Journal on Selected Areas in Communications*, 11(8):1268–1277, October 1993.

[CCI91a] CCITT. *Recommendation I.361:B-ISDN ATM Layer Sspecification.* CCITT, 1991.

[CCI91b] CCITT. *Recommendation I.362:B-ISDN ATM Adaptation Layer (AAL) Function.* CCITT, 1991.

[CCI91c] CCITT. *Recommendation I.362:B-ISDN ATM Adaptation Layer (AAL) Functional Description.* CCITT, 1991.

[CCO93] J. Chen, I. Cidon, and Y. Ofek. A local fairness algorithm for gigabit LANs/MANs with spatial reuse. *IEEE Journal on Selected Areas in Communications*, 11(8):1183–1192, October 1993.

[CDGK92] I. Cidon, J. Derby, I. Gopal, and B. Kadaba. A Critique of ATM from a Data Communications Perspective. *Journal of High Speed Networks*, 1(4):315–336, 1992.

[CGL91] M. Conti, E. Gregori, and L. Lenzini. A methodological approach to an extensive analysis of DQDB performance. *IEEE Journal on Selected Areas in Communications*, 9(1):76–87, January 1991.

[CO89a] I. Cidon and Y. Ofek. Distributed Fairness Algorithm for Local Area Networks with Concurrent Transmissions. In *Proc. of 3rd International Workshop on Distributed Algoriths*, pages 57–69, Sept. 1989.

[CO89b] I. Cidon and Y. Ofek. METARING—a ring with fairness and spatial reuse. Technical report, IBM T.J. Watson Research Center, 1989.

[DB88] D. Dykeman and W. Bux. Analysis and tuning of the FDDI media access protocol. *IEEE Journal on Selected Areas in Communications*, 6:997–1010, July 1988.

[DG92a] W. Dobosiewicz and P. Gburzyński. The very weak token protocol for fast ring networks. 1992. (Submitted for publication.).

[DG92b] W. Dobosiewicz and P. Gburzyński. The weak token protocol for high-speed ring networks. 1992. (Submitted to *IEEE INFOCOM'93*.).

[DGR91] W. Dobosiewicz, P. Gburzyński, and P. Rudnicki. Dynamic recognition of the configuration of bus networks. *Computer Communications*, 14(4):216–222, May 1991.

[DM90] P. Davids and P. Martini. Performance analysis of DQDB. In *Proceedings of IEEE IPCCC*, pages 548–555, Phoenix, AZ, 1990.

[DQD91] Distributed Queue Dual Bus Subnetwork of a Metropolitan Area Network. IEEE Std 802.6–1990, July 1991.

[Eth80] The Ethernet, A Local Area Network, Data Link Layer and Physical Layer Specifications. Digital Equipment Corporation, Intel Corporation, Xerox Corporation, Version 1.0, September 1980.

[FDD87] Fiber Distributed Data Interface (FDDI) – Token Ring Media Access Control (MAC). American National Standard for Information Systems, Doc. No. X3, 139–1987, November 1987.

[FDD90] Fiber distributed data interface, System Level Description. Digital Equipment Corporation, June 1990.

[FMO+91] E.C. Foudriat, K. Maly, C.M. Overstreet, S. Khanna, and F Paterra. A Carrier Senced Multiple Access Protocol for High Data Ring Networks. *ACM Computer Communications Review*, 21(2):59–70, April 1991.

[Gbu93] P. Gburzynski. *Protocol modeling in SMURPH.* Prentice-Hall, 1993. (In preparation.).

[Hal92] F. Halsall. *Data Comunications, Computer Networks and Open Systems.* Addison–Wesley Publishing Company, Reading, Ma., 1992.

[IEEa] IEEE (1985). *802.3 CSMA/CD Access Method and Phisical Layer Specification.* IEEE.

[IEEb] IEEE (1985). *802.5 Token Ring Access Method and Phisical Layer Specification.* IEEE.

[Jai89] R. Jain. Performance analysis of FDDI token ring networks: Effect of parameters and guidelines for setting TTRT. Technical Report DEC-TR-655, Digital Equipment Corporation, 1989.

[Kle75] L Kleinrock. *Queuing Systems.* John Wiley & Sons, Inc., 1975.

[LF82] J.O. Limb and C. Flores. Description of Fasnet – A Unidirectional local area communication network. *Bell Syst. Tech. J.*, 61:1413–1440, Sept. 1982.

[Lim90] J. Limb. A simple multiple access protocol for metropolitan area networks. In *SIGCOM*, pages 67–79, Philadelphia, PA, Sept. 1990.

[Max85] N.F. Maxemchuk. The Manhattan street network. In *Proceedings of GLOBECOM'85*, pages 255–261, 1985.

[MNW90] H.R. Mueller, M.M Nassehi, and J.W. Wong. DQMA and CRMA: New access schemes for Gbit/s LANs and MANs. In *Proceedings of INFOCOM'90*, San Francisco, CA, June 1990.

[Nas90] N.M. Nassehi. CRMA: An access scheme for high-speed LANs and MANs. In *Proceedings of SUPERCOMM/ICC'90*, Atlanta, GA, April 1990.

[Ros89] F.E. Ross. An overview of FDDI: The fiber distributed data interface. *IEEE Journal on Selected Areas in Communications*, 7(7):1043–1051, September 1989.

[RS90] R Rom and M. Sidi. *Multiple Access Protocols – Performance and Analysis*. Springer–Verlag, New York, 1990.

[SS93] O. Sharon and A. Segall. A simple scheme for slot reuse without latency for a dual bus configuration. *IEEE/ACM Transactions on Networking*, 1(1):96–104, February 1993.

[Sta87] W. Stallings. *Local Networks: An Introduction*. Macmillan, 1987.

[vA90] H.R. van As. Performance evaluation of bandwidth balancing in the dqdb mac protocol. In *Eigth Annual European Fibre Optic and Local Area Networks Conference EFOC/LAN 90*, pages 231–239, June 1990.

[vALZ92] H.R. van As, W.W. Lemppenau, and P. Zafiropulo. Performance of CRMA-II: a reservation-based fair media access protocol. In *Proceedings of EFOC-LAN 92*, Paris, France, June 1992.

[Won89] J.W. Wong. Throughput of DQDB networks under heavy load. In *7th Europ. Fibre Optic Commun. Local Area Netw. Expo.*, pages 146–151, Amsterdam, The Netherlands, June 1989.

[WOS92] H.T. Wu, Y. Ofek, and K. Sohraby. Integration of synchronous and asynchronous traffic on the metaring architecture and its analysis. *IBM Technical Report*, RC 17718, 1992.

Appendix A

Dual Chordal Bus

Section 3.1.1 shows the theoretical proof of the maximum throughput for the bus architecture. This Section shows theoretical formulas for the maximum throughput and the optimum placement of three wiring closet switches for the chordal bus architecture.

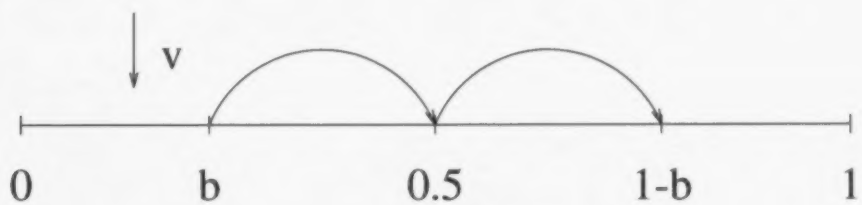


Figure A.41:

As in the case of the plain bus architecture the length of the bus is equal to 1. The dual bus consist of the infinite number of stations evenly placed on the bus. The traffic pattern is uniform, i.e., the probability that a packet is addressed to a station is equal for all stations. vdt denotes the volume sent by a station in time dt. Three wiring closet nodes are places on the bus: first in the middle of the two other in distance b from the ends of the bus (see Figure A.41).

To find the optimum placement of the wiring closet switches, point b on the bus must be such that the load in this point and in the point 0.5 (the middle of the bus) is equal 1.

The load in the section [0, b] is equal (see Section 3.1.1):

$$vx(1-x)$$

The load in the section [b, 0.5] is divided between the loads in the chordal link and section [b, 0.5] of the bus.

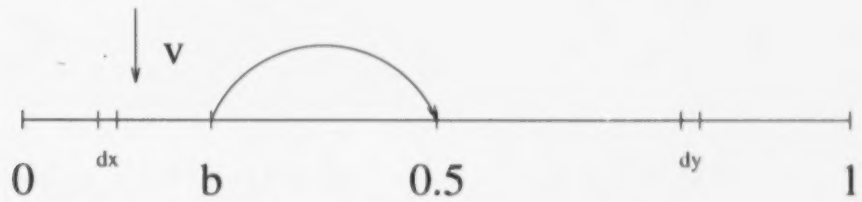


Figure A.42:

The load in the chordal link is calculated assuming, that the link conducts the traffic between section [0, b] and section [0.5, 1] (see Figure A.42). Thus, the load in the chordal link can be calculated as:

$$\int_0^b \int_{1/2}^1 v(1-x)dx\frac{dy}{1-x} = \frac{vb}{2}$$

so the load on the section [b, 0.5] of the bus is equal:

$$vx(1-x) - \frac{vb}{2}$$

To find the maximum load v and the optimum placement of the wiring closet nodes b two equations must be solved:

$$vb(1-b) = 1$$

$$v\frac{1}{2}(1-\frac{1}{2}) - \frac{vb}{2} = 1$$

The equations above are solved for the following values of v and b:

$$v = \frac{16}{(3-\sqrt{5})(1+\sqrt{5})} \approx 6.47 \qquad b = \frac{3}{4} - \frac{1}{4}\sqrt{5} \approx 0.19$$

It is easy to show that the load in the second chordal link is equal to:

$$v\int_0^{1/2} \int_{1-b}^1 dxdy = \frac{vb}{2}$$

and is identical to the load in the first chordal link.

Appendix B

Dual Chordal Ring

This Section shows theoretical formulas for the maximum throughput and the optimum number of wiring closet switches for the chordal ring architecture (see Figure 3.13).

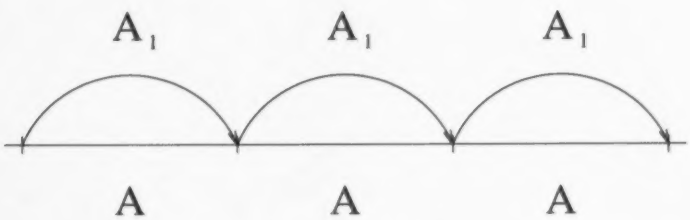


Figure B.43:

Since the chordal ring architecture is symmetric, we can assume that the load in the different sections of the ring is also symmetric. There are two kinds of sections of the ring: the chordal link section, denoted A_1 (see Figure B.43) and regular link sections, denoted A. All regular stations are connected to regular link section; none of them is connected to the chordal link section. The connection between the chordal link section and the regular link section is done by the wiring closet switch. The following assumption are made:

- The length of the rings is equal to 1.

- The dual ring consists of an infinite number of stations evenly placed on the ring.
- The traffic pattern is uniform, i.e., the probability that a packet is addressed to a station is equal for all stations. Likewise, the traffic volume sent by each station is the same: *vdt* denotes the volume sent by a station in time *dt*.
- We also assume that n wiring closet nodes are places on the ring, at intervals of length $1/n$.

To find the optimal number of wiring closet switches and the maximum load of the chordal ring architecture, we must find functions describing load in sections A and A_1. The load in any of these sections cannot exceed 1 (i.e., the normalized capacity of each link).

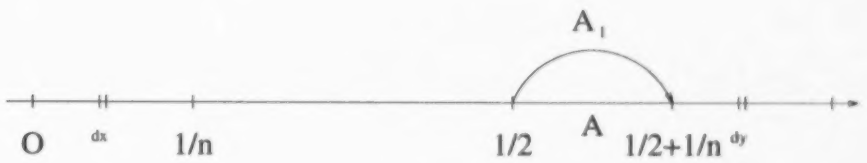


Figure B.44:

To find the function describing the load in the chordal link section A_1 consider the situation presented in Figure B.44. The load in the link A_1 can be viewed as the following sum:

$$\sum_{i=1}^{\frac{n}{2}-2} \frac{v}{n^2} i + \sum_{i=1}^{\frac{n}{2}-1} \frac{v}{2n^2} = \frac{v}{8}\left(\frac{n-2}{n}\right)^2$$

The first sum describes the sum of loads from the entire section placed to the left of the chordal link to another entire section placed to the right of the chordal link and described as a sum of:

$$\int_0^{\frac{1}{n}} \int_0^{\frac{1}{n}} \frac{1}{2} v dx \frac{dy}{\frac{1}{2}} = \frac{v}{n^2}$$

The second sum describes the sum of loads from a part of a section placed to the left of a chordal ring to a part of a section placed to the right of the chordal link and

described as a sum of:

$$\int_0^{\frac{1}{n}} \int_{\frac{1}{2}}^{\frac{1}{2}+x} \frac{1}{2} v dx \frac{dy}{\frac{1}{2}} = \frac{v}{2n^2}$$

The load in the section A can be viewed as a local and non-local one. The local load at station $s \in [0, 1/n]$ can be calculated as the difference between the load caused by stations in section [0, 1/n] minus the amount of the load caused by packets sent and received in the same section [0, 1/n]. Thus it is equal to:

$$\int_{\frac{1}{n}}^{s} \frac{v}{2} dx - \int_{\frac{1}{2}}^{s} \int_{\frac{1}{2}}^{s} \frac{v}{2} dx \frac{dy}{\frac{1}{2}} = \frac{sv}{2}(1-s)$$

The non-local load is caused by packets addressed to stations in the section [0, 1/n], thus the non-local load at station s is equal to[1]:

$$\int_s^{\frac{1}{n}} \left(\frac{1}{2} - x\right) v dx = v\left(\frac{1}{2n} - \frac{s}{2} - \frac{1}{2n^2} + \frac{s^2}{2}\right)$$

The the load in the section A is thus equal to:

$$v\left(\frac{1}{2n} - \frac{1}{2n^2}\right)$$

and is constant in the whole section.

To find the maximum load v and the optimum number of the wiring closet nodes n two equations must be solved:

$$\frac{v}{8}\left(\frac{n-2}{n}\right)^2 = 1$$

$$v\left(\frac{1}{2n} - \frac{1}{2n^2}\right) = 1$$

The equations above are solved for the following values of v and n:

$$v = 16 \qquad n = 4 + 2\sqrt{2} \approx 6.82$$

[1]Notice, that the maximum distance between two stations is equal to 1/2.